*UNIVERSITE TOULOUSE III - PAUL SABATIER*
*ECOLE DOCTORALE DES SCIENCES DE L'UNIVERS*
*DE L'ENVIRONNEMENT ET DE L'ESPACE*

*UNIVERSIDADE DE ÉVORA*
*DEPARTAMENTO DE FÍSICA*

***THESE EN CO-TUTELLE***

pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITE TOULOUSE III et DE L'UNIVERSITE D' ÉVORA**

***Discipline :*** **Physique de l'Atmosphère**

présentée et soutenue

par

**Mme. Margarida Belo Pereira**

le Mai 2006

***Titre :***

# ESTIMATION AND STUDY OF FORECAST ERROR COVARIANCES USING AN ENSEMBLE METHOD IN A GLOBAL NWP MODEL

## Utilisation d'une méthode d'ensemble pour estimer et étudier les covariances d'erreur dans un modèle météorologique global

## Utilização de um método de ensemble para estimar e estudar as estatísticas dos erros de um modelo de previsão numérica do tempo




***JURY:***

| | | |
|---|---|---|
| M. | **António Heitor Reis** | Président |
| M. | **João Corte Real** | Directeur de thèse |
| M. | **Patrick Mascart** | Directeur de thèse |
| M. | **Carlos Pires** | Rapporteur |
| M. | **François-Xavier Le Dimet** | Rapporteur |
| M. | **Loïk Berre** | Examinateur |
| M. | **Serge Chauzy** | Examinateur |
| M. | **Alfredo Rocha** | Examinateur |

METEO-FRANCE/CNRM/GMAP
42 Avenue Coriolis - 31057 Toulouse, France

Instituto de Meteorologia/DOV/CMAR
Rua C ao Aeroporto - 1800 Lisboa, Portugal

# Acknowledgements

I wish to thank all the people who have made this study possible. This work was mainly funded by the ALAdin Training NETwork project, ALATNET, coordinated by Jean-François Geleyn, contract no. HPRN-CT-1999-00057, a European Union initiative under the Fifth Framework Programme.

During my stay in France I was employed by Mètèo-France at "Centre National Recherches Mtorologiques (CNRM)", in the "Groupe Modelisation en Assimilation et Prevision (GMAP)" section which is headed by Dr. François Bouttier and was headed by Jean-François Geleyn during the first stage of this work.

In Portugal I was employed by the Portuguese Meteolorological Service, under the "Previsão Numérica, Arquivo e Processamento" section (headed by Maria José Monteiro), which belongs to the "Departamento de Vigilância Meteorológica" directed by Teresa Abrantes.

The Ph.D. study has been conducted under a "co-tutelle" agreement between the french university Paul Sabatier, Toulouse III, and the Portuguese University of Évora.

In particular, I wish to thank to Dr. Loïk Berre, for carefully supervising this work, for his help, for the precious discussions we had and for encouraging me to write.

I am grateful to Prof. Patrick Mascart for coordinating the Ph.D. in France and for his encouragement during the study.

I wish to thank the GMAP team and the ALATNET/ALADIN scientific community that I had the pleasure of meeting, for the many interesting and helpful discussions. In particular, I wish to thank to Dominique Giard, Jean-François Geleyn, François. Bouyssel, Claude Fischer, Ryad Elkhatib, Eric Escaliere and Jean Maziejewski. To Dr. François Bouttier and Dr. Gerald Desroziers I am also thankful for their comments and suggestions which were helpful for the writing of the manuscript.

Special thanks to Goska Szczech, Andrè Simon, Gianpaolo Balsamo, Cornel Soci, Doina Banciu, Marian Jurasek, Mateja Irsic and Stjepan Ivatek-Sahdan, with whom I also spent nice time in Toulouse.

I am also grateful to several people in Portugal who have also contributed to this work. I wish to thank to Prof. João Corte Real who coordinated the Ph.D. in Portugal and always encourage me to do this study.

I wish to thank to my colleagues of the Portuguese Meteorological Service, who encourage me during this study, in particular to my friends Nuno Moreira and Isabel Soares.

I am also thankful for the support of my family and friends, in particular to, José Pedro, Ana Teresa, Matilde Vieira, Cristina Mendes, José and Paula. Moreover, I would like to thank to my parents for showing me that someone should always fight to achieve their goals. I am deeply grateful to my grandparents Joaquim and Ana Maria, for always being there in the important moments. Finally, I am thankful to my fiancé Filipe, for his precious support throughout the last three years.

*dedicated to my grandparents,*

# Synthesis

The production of an accurate analysis is one important goal of modern NWP centers, where sophisticated data assimilation techniques (such as variational methods) have been implemented. However, this task is not straightforward, since the real state of the atmosphere is never exactly known.

One of the main difficulties in data assimilation is caused by the fact that the degrees of freedom of the modern NWP models ($\sim 10^7$) are larger than the number of independent available observations ($\sim 10^5$). Moreover, the distribution of the observation network is not uniform in space and in time. For these reasons, it is not enough to perform a spatial interpolation of observations into a regular grid. A prior information is needed in order to solve the undetermined analysis problem. In other words, it is necessary to have a first guess about the atmospheric state at all grid points. In modern data assimilation schemes, this first guess (or background) is provided by a short range forecast (from a previous analysis cycle). Hence, the analysis field results from a combination of observations and a background field.

The relative weights given to observations and to the background depend on specified observation and background error covariance matrices (which are usually noted $\mathbf{R}$ and $\mathbf{B}$, respectively). More precisely, the filtering and interpolation of the observations are mostly determined by the expected magnitudes of the errors at each location (variances) and by the expected correlations between the errors at different locations. The variances are the diagonal terms of the covariance matrices and the correlations correspond to the off-diagonal terms. Therefore, the quality of the analysis strongly depends on the accuracy of the observation and background error covariance estimates.

The $\mathbf{B}$ matrix is also important because it contains the balance relationships between mass and wind fields (e.g., geostrophy in the mid-latitudes): this allows the mass analysis to be influenced by wind observations, and vice-versa.

The main goal of the present work is to study and improve the estimate of the background error covariance matrix. In order to achieve this, an ensemble approach has been applied to the Arpège global model in its non stretched version (i.e., with a uniform resolution on the globe). This technique is based on an ensemble of independent perturbed analysis experiments, and it was first proposed by Houtekamer *et al.* (1996).

In the ensemble approach, for each experiment (member) and for each analysis cycle, all observations are perturbed by adding independent random numbers (which have a Gaussian distribution with mean zero and variance equal to the prescribed variance of the observation error). Some background perturbations can be also constructed, and these two perturbations (of the observations and of the background) will create a perturbed analysis for each member. From each perturbed analysis, a 6h integration of the numerical model is performed. In this way, an ensemble of perturbed analyses will evolve into an ensemble of perturbed backgrounds. For each ensemble member, it is possible to use either the same forecast model or different models. Moreover, the ensemble experiments can be carried out for several consecutive cycles, for instance during a one or two month period. It is assumed that the differences between the perturbed background fields of the different members represent the background errors. Thus, the background error covariances are estimated from these background differences. In the same way, it is possible to estimate the analysis error covariances from the ensemble analysis differences.

In the first stage of the work, the ensemble method has been implemented using a perfect-model framework. The resulting covariance estimates have been compared with those that were previously operational in the Arpège 4D-Var, by performing both diagnostic and impact studies (Belo Pereira and Berre, 2006).

In particular, the global (i.e. spatially averaged) ensemble covariances have been compared with those of the NMC method. This NMC technique was used only to specify the global

covariances. This method relies on the assumption that the background error can be estimated by differences between forecasts of different ranges (for instance 12 and 36 hours), but valid at the same time. This technique is known as the NMC method, because it was first implemented in the U.S. National Meteorological Center (now renamed National Center for Environmental Prediction) (Parrish and Derber 1992). The NMC method has the advantage of being very easy to implement. This probably explains that this method was (or still is) operational in many NWP centers.

It was found that, globally, the ensemble approach emphasizes the relative contributions of the small scales to the background errors. Consequently, the global length scales (which describe how fast the correlation decreases with distance) are smaller in the ensemble method than in the NMC method. These results are in accordance with those described by Fisher (2003), where these differences were considered to be caused by the involvement of longer forecast ranges in the NMC method.

Nevertheless, the analysis error estimated by the NMC method, strongly depends on the analysis increment (difference between analysis and background), and it was found that its spectrum is much larger scale than the analysis dispersion spectrum from the ensemble method. In other words, the large scale contributions are enhanced in the analysis increment spectrum, when compared with the spectrum of the analysis error estimated by the ensemble method. Moreover, this result is consistent with the expected filtering properties of the **B** matrix in the analysis equation. This suggests that the different representations of the analysis step, in the simulation of the error evolution, also explain an important part of the differences between the two methods.

The local (i.e. spatially varying) ensemble variances have been compared with those provided by the NMC method, and with the local operational variances (which correspond to a static 3D map of variances, that reflects some typical data density contrasts).

Previous studies (Bouttier (1994) and McNally (2000)) suggest that the analysis error variances estimated by the NMC method are underestimated in the data-poor areas, and overestimated in regions where the observations are dense and have a high quality. For these reasons, in the Arpège 4D-Var, the NMC method is not used to provide the spatial variations of the background error variance. The examination of the local covariances, and the comparison between the analysis errors and the background errors estimated by the NMC method, are consistent with the referred studies.

In contrast with this, the local ensemble background error variances capture some interesting features, such as the contrasts between data-rich and data-sparse regions, and also the areas of large atmospheric variability (the ITCZ, for instance). Moreover, the analysis dispersion provided by the ensemble approach appears to give some relevant information about the analysis error, for instance that concerning the influence of data density.

The local correlation structures were also examined. For this purpose, an economical method was developed to estimate the local correlation length scales. The corresponding estimates provided by the ensemble method show that the length scale is largest in the tropics and smallest in the data rich regions, as expected from other studies (e.g. Lindzen and Fox-Rabinovitz (1989), Ingleby (2001), Bouttier (1994)). In addition, these estimates from the NMC method differ from the ensemble ones, mainly for temperature. In particular, the results indicate that the overestimation of the length scales by the NMC method is more pronounced over data rich areas, such as North America. These differences suggest that the expected data density effects on the correlation structures are not well represented by the NMC method.

Moreover, this local length scale diagnosis also allows to examine the correlation anisotropies. This information about heterogeneities and anisotropies can be useful to model spatially varying correlations in the **B** matrix, e.g. by wavelet techniques (Fisher (2003), Deckmyn and Berre (2005)).

Some impact studies were also performed with the Arpège 4D-Var. The use of the ensemble global covariances (compared with the NMC global covariances) had a general positive impact

on the forecast quality. This positive impact was enhanced, when the ensemble local standard deviation maps (instead of the former operational ones) were used in the minimization and quality control stages. For the above reasons, the ensemble statistics became operational at Météo France in January 2004.

However, one could expect that the perfect model assumption causes an underestimation of the background error variances. Therefore, the impact of simulating the model error in the ensemble approach has also been studied. The model error is simulated by performing an ensemble of experiments using different resolutions and different values for certain tuning parameters of some physical parametrizations. The resulting covariances are compared with those provided by the ensemble approach in a perfect-model context. Furthermore, the impact of the corresponding covariance estimates on the forecast quality was also studied.

The results indicate that the resolution is the most important contribution for small scale variables (e.g., vorticity and divergence). The perturbations of the physical parametrizations remain important in the tropics for vorticity and divergence, and also for temperature in general. This is more obvious in lower and middle troposphere in the extra-tropics and in upper levels over the tropics.

Moreover, the effect of the model error simulation in the ensemble approach is to increase the magnitude of the estimated background error. For large scale variables (temperature and surface pressure), the increase in magnitude of the background errors is associated to the enhancement of the synoptic scale contributions. On the other hand, for the other variables (vorticity, divergence, and humidity), the simulation of the model error appears to emphasize the relative contributions of the mesoscale phenomena to the background errors, mainly due to the simulation of the model error related to resolution. This causes a reduction of the coupling between mass and wind variables.

The referred reduction of the mass/wind coupling has a positive impact on the forecast quality of geopotential over the tropics and in some stratospheric extra-tropical areas. On the other hand, the reduction of mass/wind balance leads to a degradation of the forecasts of geopotential and wind in some tropospheric extra-tropical regions. This suggests that the mass/wind coupling is smaller in the tropics and in the stratosphere (as could be expected), and that the relaxation of this coupling as a function of latitude and height should be strengthened in the $\mathbf{B}$ matrix.

It was also found that there is a better agreement between the vertical profiles of the statistics of the innovations (differences between observations and background) and those from the ensemble approach, when the model error is simulated.

Finally, the effect of neglecting the observation errors in the ensemble method has also been studied. For this purpose, the observation perturbations have been set to zero. This means that the ensemble simulations have been performed assuming that the background error is caused mainly by the model error (either due to uncertainties in physical parametrizations and resolution, or only due to the uncertainties in parametrizations). The resulting covariances have been compared with those in a perfect-model context, and with those in which both the observation and model errors are simulated.

Globally, it was found that the assumption that observations are perfect leads to a decrease of the background error variance, mostly due to a reduction of the large scale contributions. This effect is consistent with the filtering properties of the analysis equation. It is more pronounced in upper troposphere and for surface pressure.

In addition, it was found that the magnitude of the background error estimated by the ensemble approach is smaller when it is assumed that observations are perfect rather than when the model is assumed to be perfect. There are two possible interpretations of this result. The first one is that the uncertainties related to the observations have a larger contribution to the background error than the model error. The second one is that some important sources of model

error have been underestimated.

The sensitivity of the ensemble estimates to the model resolution has also been examined. For this purpose, two background error statistics have been derived from two similar ensemble perfect model experiments, which were built using the same model at two different resolutions. It was found that, when the model truncation increases, the magnitude of the background dispersion decreases in the large scales and increases in the small scales. At first sight, the increase of small scale variance may be seen as a paradox. However, it is important to note that as the model truncation decreases, its variability at smaller scales will be smaller, because these scales are either not represented at all by the model, or dissipated by the horizontal diffusion. Moreover, the small scale structures that are observed but not represented by the model, due to its limited resolution, are usually considered and treated as a part of the observation representativeness errors (rather than as a part of the background errors).

For the above reasons, the determination of the optimum amount of small scale variance, for the specification of the **B** matrix in data assimilation, remains an open issue.

On the whole, the results indicate that the ensemble method provides some relevant and useful information on the time-averaged background error covariances, both globally and locally. This suggests that one should also try to extract some flow-dependent information from the ensemble dispersion. The use of this analysis ensemble approach for high resolution limited area models is another natural application, as showed by recent studies with the Aladin model (Ştefănescu, Berre and Belo Pereira 2006, Berre, Ştefănescu and Belo Pereira 2006).

# Synthèse

La production d'une analyse précise est un objectif important des centres modernes de prévision numérique du temps, où des techniques sophistiquées d'assimilation de données (telles que les méthodes variationnelles) ont été implémentées. Cependant, cette tâche n'est pas triviale, parce que l'état réel de l'atmosphère n'est jamais connu de façon exacte.

L'une des difficultés principales en assimilation de données est liée au fait que les degrés de liberté des modèles modernes de prévision numérique ($\sim 10^7$) sont plus nombreux que le nombre d'observations indépendantes disponibles ($\sim 10^5$). De plus, la distribution du réseau d'observation n'est par uniforme dans l'espace et dans le temps. Pour ces raisons, il n'est pas suffisant de réaliser une interpolation spatiale des observations sur une grille régulière. Une information préalable est requise afin de résoudre le problème sous-déterminé qu'est l'analyse. En d'autres termes, il est nécessaire de disposer d'une ébauche relative à l'état atmosphérique sur tous les points de grille. Dans les schémas modernes d'assimilation de données, cette ébauche est fournie par une prévision à courte échéance (d'un cycle d'analyse précédent). En conséquence, l'analyse résulte d'une combinaison des observations et d'une ébauche.

Les poids relatifs donnés aux observations et à l'ébauche dépendent des matrices spécifiées de covariance des erreurs d'observation et de l'ébauche (qui sont habituellement notées **R** et **B**, respectivement). Plus précisément, le filtrage et l'interpolation des observations sont principalement déterminés par les amplitudes attendues des erreurs en chaque point (variances) et par les corrélations attendues entre les erreurs en des points distincts. Les variances sont les termes diagonaux des matrices de covariance, et les corrélations correspondent aux termes non diagonaux. En conséquence, la qualité de l'analyse dépend fortement de la précision des estimations des covariances d'erreur des observations et de l'ébauche.

La matrice **B** est également importante parce qu'elle contient les relations d'équilibre entre les champs de masse et de vent (par exemple, le géostrophisme aux moyennes latitudes): cela permet à l'analyse du champ de masse d'être influencée par les observations de vent, et vice-versa.

Le but principal du travail présent est d'étudier et d'améliorer l'estimation de la matrice de covariance d'erreur d'ébauche. A cette fin, une méthode d'ensemble a été appliquée au modèle global Arpège dans sa version non étirée (c'est-à-dire avec une résolution uniforme sur le globe). Cette technique est basée sur un ensemble d'expériences d'analyses perturbées de façon indépendante. Elle a été proposée initialement par Houtekamer et al (1996).

Dans la méthode ensembliste, pour chaque expérience (membre) et pour chaque cycle d'analyse, toutes les observations sont perturbées en ajoutant des nombres aléatoires indépendants (qui suivent une distribution gaussienne de moyenne nulle, et dont la variance est égale à la variance spécifiée de l'erreur d'observation). Des perturbations de l'ébauche peuvent aussi être construites, et ces deux perturbations (des observations et de l'ébauche) vont créer une analyse perturbée pour chaque membre. A partir de chaque analyse perturbée, une intégration sur 6h du modèle numérique est réalisée. De cette façon, un ensemble d'analyses perturbées va évoluer en un ensemble d'ébauches perturbées. Pour chaque membre de l'ensemble, il est possible d'utiliser soit le même modèle de prévision ou différents modèles. De plus, les expériences ensemblistes peuvent être menées sur plusieurs cycles consécutifs, par exemple durant une période d'un ou deux mois. Les différences entre les champs perturbés d'ébauche des différents membres sont supposées représenter les erreurs de prévision. Ainsi, les covariances d'erreur d'ébauche sont estimées à partir des différences d'ébauche. De la même façon, il est possible d'estimer les covariances d'erreur d'analyse à partir des différences d'analyse de l'ensemble.

Au cours de la première étape de ce travail, la méthode ensembliste a été mise en oeuvre dans un contexte de modèle parfait. Les estimations de covariance résultantes ont été comparées à celles qui étaient opérationnelles précédemment dans le 4D-Var Arpège, en réalisant des études

diagnostiques et d'impact (Belo Pereira et Berre 2006).

En particulier, les covariances ensemblistes globales (i.e. moyennées spatialement) ont été comparées à celles de la méthode du NMC. Cette technique du NMC n'était utilisée que pour spécifier les covariances globales. La méthode du NMC repose sur l'hypothèse que l'erreur d'ébauche peut être estimée par les différences entre des prévisions dont les échéances varient (par exemple 12 et 36 heures), mais qui sont valables pour la même date. Cette technique est connu sous le nom de "méthode du NMC", parce qu'elle a été mise en oeuvre pour la première fois au Centre Météorologique National des Etats-Unis (qui s'appelle maintenant le Centre National pour la Prévision Environnementale) (Parrish et Derber 1992). La méthode du NMC a l'avantage d'être très facile à mettre en oeuvre. Cela explique probablement que cette méthode était (ou est encore) opérationnelle dans beaucoup de centres de prévision numérique.

Il est apparu que, globalement, la méthode ensembliste met davantage l'accent sur les contributions relatives des petites échelles aux erreurs d'ébauche. En conséquence, les portées globales (qui décrivent la vitesse de décroissance de la corrélation avec la distance) sont plus petites dans la méthode ensembliste que dans la méthode du NMC. Ces résultats sont en accord avec ceux décrits par Fisher (2003), pour lesquels ces différences étaient considérées comme étant causées par les échéances plus longues de la méthode du NMC.

Néanmoins, l'erreur d'analyse estimée par la méthode du NMC dépend fortement de l'incrément d'analyse (qui est la différence entre l'analyse et l'ébauche), et il est apparu que son spectre est nettement plus grande échelle que le spectre de dispersion de l'analyse estimé par la méthode ensembliste. De plus, ce résultat est cohérent avec les propriétés de filtrage attendues de la matrice $\mathbf{B}$ au niveau des équations de l'état d'analyse et de l'incrément d'analyse. Cela suggère que les différentes représentations de l'étape d'analyse, dans la simulation de l'évolution des erreurs, expliquent également une part importante des différences entre les deux méthodes.

Les variances ensemblistes locales (i.e. variables spatialement) ont été comparées à celles fournies par la méthode du NMC, ainsi qu'avec les variances opérationnelles locales (qui correspondent à une carte 3D statique de variance, qui reflète certains contrastes typiques de densité de données).

Des études précédentes (Bouttier (1994) et McNally (2000)) suggèrent que les variances d'erreur d'analyse estimées par la méthode du NMC sont sous-estimées dans les régions pauvres en observations, et surestimées là où les observations sont denses et de bonne qualité. Pour ces raisons, dans le 4D-Var Arpège, la méthode du NMC n'est pas utilisée pour fournir les variations spatiales de la variance d'erreur d'ébauche. L'examen des covariances locales, et la comparaison entre les erreurs d'analyse et d'ébauche estimées par la méthode du NMC, sont cohérentes avec les études précitées.

Par contre, les variances ensemblistes d'erreur d'ébauche permettent de représenter des caractéristiques intéressantes, telles que les contrastes liés à la densité des observations, ainsi que les régions où la variabilité atmosphérique est forte (la ZCIT par exemple). De plus, la dispersion d'analyse fournie par l'approche ensembliste s'avère donner une information pertinente sur l'erreur d'analyse, par exemple concernant l'influence de la densité des observations.

Les structures locales de corrélation ont également été examinées. A cette fin, une méthode économique a été développée pour estimer les portées locales de corrélation. Les estimations correspondantes fournies par la méthode ensembliste montrent que la portée est maximale dans les régions tropicales, et minimale dans les régions riches en observations, en accord avec d'autres études (Lindzen et Fox-Rabinovitz (1989), Ingleby (2001), Bouttier (1994)). De plus, les estimations de la méthode du NMC diffèrent des estimations ensemblistes, principalement pour la température. En particulier, les résultats indiquent que la surestimation des portées par la méthode du NMC est plus prononcée sur les régions riches en observations, telles que l'Amérique du Nord. Ces différences suggèrent que les effets attendus de la densité des observations, sur les structures de corrélation, ne sont pas bien représentés par la méthode du NMC.

De plus, ce diagnostic de portée locale permet aussi d'examiner les anisotropies de corrélation. Cette information sur les hétérogénéités et les anisotropies peut être utile pour modéliser des

corrélations variables spatialement dans la matrice **B**, par exemple à l'aide de techniques ondelette (Fisher (2003), Deckmyn et Berre (2005)).

Des études d'impact ont également été réalisées avec le 4D-Var Arpège. L'utilisation des covariances globales ensemblistes (comparées aux covariances globales NMC) ont eu un impact positif général sur la qualité des prévisions. Cet impact positif a été renforcé, quand les cartes d'écart type local ensembliste ont été utilisées (à la place de celles qui étaient opérationnelles auparavant) pour la minimisation et le contrôle de qualité. Pour ces raisons, les statistiques ensemblistes sont devenues opérationnelles à Météo France en janvier 2004.

Cependant, on pourrait s'attendre à ce que l'hypothèse de modèle parfait cause une sous-estimation des variances d'erreur d'ébauche. En conséquence, l'impact d'une simulation de l'erreur de modèle dans l'approche ensembliste a été étudié. L'erreur de modèle est simulée en réalisant un ensemble d'expériences qui utilisent des résolutions différentes et des valeurs différentes pour des paramètres de réglage de certaines paramétrisations physiques. Les covariances résultantes ont éte comparées à celles fournies par l'approche ensembliste dans un contexte de modèle parfait. Par ailleurs, l'impact des estimations correspondantes de covariance sur la qualité des prévisions a également été étudié.

Les résultats indiquent que la résolution est la contribution la plus importante pour les variables de petite échelle (tourbillon et divergence). Les perturbations des paramétrisations physiques restent importantes sous les tropiques pour le tourbillon et la divergence, ainsi que pour la température en général. Cela est plus net dans la basse et moyenne troposphère aux moyennes latitudes et pour les niveaux élevés sous les tropiques.

De plus, l'effet de la simulation d'erreur dans l'approche ensembliste est d'augmenter l'amplitude de l'erreur d'ébauche estimée. Pour les variables de grande échelle (température et pression de surface), l'augmentation d'amplitude des erreurs d'ébauche est associée au renforcement des contributions d'échelle synoptique. D'un autre côté, pour les autres variables (tourbillon, divergence, et humidité), la simulation de l'erreur de modèle s'avère renforcer les contributions relatives des phénomènes mésoéchelle aux erreurs d'ébauche, principalement du fait de la simulation de l'erreur de modèle liée à la résolution. Cela cause une réduction du couplage entre les variables de masse et de vent.

Cette réduction du couplage masse/vent a un impact positif sur la qualité de prévision du géopotentiel sous les tropiques, ainsi que dans certains régions stratosphériques extratropicales. D'un autre côté, la réduction de l'équilibre masse/vent conduit à une dégradation des prévisions de géopotentiel et de vent dans certaines régions troposphériques extratropicales. Cela suggère que le couplage masse/vent est plus faible sous les tropiques et dans la stratosphère (comme on pouvait s'y attendre), et que la relaxation de ce couplage en fonction de la latitude et de l'altitude devrait être renforcée dans la matrice **B**.

Il est également apparu qu'il y a un meilleur accord entre les profils verticaux des statistiques des innovations (qui sont les carts entre les observations et l'ébauche) et ceux de la méthode ensembliste, lorsque l'erreur de modèle est simulée.

Finalement, l'effet de la non prise en compte des erreurs d'observation dans la méthode ensembliste a également été étudié. A cette fin, les perturbations d'observation ont été mises à zéro. Cela signifie que les simulations ensemblistes ont été réalisées en supposant que l'erreur d'ébauche est principalement due à l'erreur de modèle (soit en raison des incertitudes liées aux paramétrisations physiques et à la résolution, soit seulement en raison des incertitudes liées aux paramétrisations). Les covariances résultantes ont été comparées à celles du contexte de modèle parfait, ainsi qu'à celles pour lesquelles les erreurs d'observation et de modèle sont toutes deux simulées.

Globalement, il est apparu que l'hypothèse que les observations sont parfaites conduit à une baisse de la variance d'erreur d'ébauche, principalement en raison d'une réduction des contributions de grande échelle. Cet effet est cohérent avec les propriétés de filtrage de l'équation

d'analyse. Cela est plus prononcé dans la haute troposphère et pour la pression de surface.

De plus, il est apparu que l'amplitude de l'erreur d'ébauche estimée par l'approche ensembliste est plus petite lorsque les observations sont supposées parfaites, que quand le modèle est supposé parfait. Deux interprétations sont possibles quant à ce résultat. La première est que les incertitudes des observations ont une contribution plus grande à l'erreur d'ébauche que l'erreur de modèle. La seconde interprétation est que des sources importantes d'erreur de modèle ont été sous-estimées.

La sensibilité des estimations ensemblistes à la résolution du modèle a également été examinée. Dans cette perspective, deux statistiques d'erreur d'ébauche ont été obtenues à partir de deux expériences ensemblistes sembables dans un contexte de modèle parfait. Ces deux statistiques ont été construites en utilisant le même modèle à deux résolutions différentes. Il est apparu que, lorsque la résolution du modèle augmente, l'amplitude de la dispersion d'ébauche décroît dans les grandes échelles et augmente dans les petites échelles. A première vue, l'augmentation de la variance dans les petites échelles peut sembler être un paradoxe. Cependant, lorsque la troncature du modèle augmente, des échelles plus petites peuvent être simulées. De plus, l'amortissement des ondes plus courtes dû à la diffusion numérique est réduit. En d'autres termes, davantage de composantes de la variabilité atmosphérique sont représentées par le modèle. La détermination de la quantité optimale de variance de petite échelle reste ensuite une question ouverte, pour la spécification de la matrice **B** en assimilation de données. Cela est lié au fait que les structures de petite échelle qui sont observées mais non représentées par le modèle, en raison de sa résolution limitée, sont habituellement traitées comme une partie des erreurs de représentativité des observations (plutôt que comme une partie des erreurs d'ébauche).

Dans l'ensemble, les résultats indiquent que la méthode ensembliste fournit des informations pertinentes et utiles sur les covariances moyennées temporellement de l'erreur d'ébauche, à la fois globalement et localement. Cela suggère que l'on devrait également essayer d'extraire une information dépendante de l'écoulement à partir de la dispersion de l'ensemble. L'utilisation d'une approche ensembliste pour des modèles à aire limitée et à haute résolution est une autre application naturelle, comme cela a été montré par des études récentes avec le modèle Aladin (Ştefănescu, Berre et Belo Pereira 2006, Berre, Ştefănescu et Belo Pereira 2006).

# Resumo longo

Um dos objectivos dos maiores Serviços Meteorológicos é produzir uma análise de boa qualidade. Para este fim são utilizadas técnicas sofisticadas de assimilação de dados, como, por exemplo, métodos variacionais.

Uma das principais dificuldades na assimilação de dados deve-se ao facto do número de graus de liberdade dos modelos de Previsão Numérica do Tempo (PNT) ($\sim 10^7$) ser superior ao número de observações independentes disponíveis ($\sim 10^5$). Além disso, as observações sobre o Globo não se encontram distribuídas uniformemente no espaço e no tempo.

Por estas razões, não é suficiente interpolar espacialmente as observações para a malha regular dos modelos. É necessário ter uma estimativa inicial do estado da atmosfera, de modo a resolver um problema indeterminado. Nas técnicas modernas de assimilação de dados, esta estimativa inicial, conhecida como *estimativa a priori* e como *background* na literatura anglo-saxónica, é fornecida por uma previsão a curto prazo (geralmente 6 horas). No presente documento esta previsão é designada simplesmente por *previsão curta*. Assim, o campo da análise resulta da combinação das observações e da *previsão curta*.

A ponderação atribuída às observações e à *previsão curta* depende das estatísticas (matrizes de autocovariância, também designadas de covariância) dos erros destas duas fontes de informação. Por outras palavras, a forma como a informação proveniente das observações é utilizada para modificar o campo da *estimativa a priori* nos pontos da malha do modelo depende das variâncias e das correlações entre os erros da *previsão curta* em diferentes pontos. As variâncias reflectem a magnitude dos erros e correspondem aos termos diagonais das matrizes de covariância. Os termos não diagonais das matrizes de covariância contêm as correlações. Contudo, como o estado real da atmosfera nunca é conhecido exactamente, as matrizes de covariância não podem ser exactamente conhecidas, podem apenas ser estimadas.

A matriz de covariância dos erros da *estimativa a priori*, conhecida como matriz **B**, é também importante porque contém informação sobre as relações entre os campos do vento e da massa (por exemplo, o equilíbrio quasi-geostrófico nas latitudes médias). Estas relações permitem que a informação proveniente das observações de uma variável possa ser utilizada por outra variável.

Pelas razões mencionadas a qualidade da análise depende fortemente da exactidão das estimativas da matriz **B**. O método de NMC é uma das técnicas utilizadas para estimar as estatísticas dos erros da *previsão curta* no contexto da previsão numérica. Este método aceita a hipótese que os erros da *previsão curta* podem ser estimados pelas diferenças entre previsões do mesmo modelo com alcances diferentes (por exemplo 12 e 36 horas), mas válidas no mesmo instante. Esta técnica é conhecida como método de NMC, porque foi implementada pela primeira vez no Centro Meteorológico Nacional dos Estados Unidos (actualmente chamado Centro Nacional para a Previsão Ambiental) (Parrish e Derber 1992).

O método de NMC tem a vantagem de ser muito fácil de aplicar. Provavelmente, este facto explica, em parte, a sua utilização em regime operacional, desde 1992. No entanto, estudos precedentes (Bouttier (1994) e McNally (2000)) sugerem que o método de NMC tende a subestimar a variância (magnitude) dos erros de análise nas regiões com poucas observações e a sobrestima-la nas regiões onde a densidade de observações é elevada e a sua qualidade é boa. Por este motivo, na assimilac cão do modelo Arpège, as estimativas provenientes do método de NMC nunca foram utilizadas para descrever as variações geográficas dos erros da *previsão curta*. Este método é utilizado apenas para fornecer a média global das covariâncias (em cada nível do modelo).

No presente trabalho, um método de *ensemble* foi aplicado ao modelo global Arpège na sua versão não-estirada (isto é, com uma resolução uniforme sobre o globo), com o objectivo de estudar e melhorar a estimativa da matriz **B**. O método de *ensemble* utiliza um conjunto de experiências independentes de análise e foi proposto pela primeira vez por Houtekamer *et al.*

(1996).

No método de *ensemble*, para cada experiência (membro) e para cada ciclo de análise, todas as observações são perturbadas, adicionando números aleatórios independentes (que têm uma distribuição gaussiana com um valor médio igual a zero e variância igual à variância dos erros das observações). É também possível gerar perturbações da *estimativa a priori*, de modo a simular os erros do modelo. Os erros do modelo podem também ser simulados utilizando modelos diferentes para cada membro do *ensemble*. A partir dos campos perturbados das observações e das *estimativas a priori* obtém-se uma análise perturbada para cada membro do *ensemble*. Através da integração (durante 6 horas) do modelo numérico, com base numa análise perturbada, obtém-se uma *previsão curta* perturbada. Estas experiências podem ser realizadas para diversos ciclos de análise consecutivos, durante um período de um ou dois meses, por exemplo. Finalmente, aceita-se que as diferenças entre os campos das *previsões curtas* perturbadas de membros diferentes representam os erros da *previsão curta* (ou da *estimativa a priori*). Desta forma é possível estimar as covariâncias dos erros da *previsão curta* a partir das diferenças entre os campos de *previsões curtas* de diferentes membros do *ensemble*. De forma análoga, é possível estimar as covariâncias dos erros da análise, a partir das diferenças entre campos de análises de diferentes membros do *ensemble*.

Inicialmente, o método de *ensemble* foi aplicado assumindo que a contribuição dos erros do modelo para os erros da *previsão curta* é desprezável. Ou seja, supõe-se que as principais fontes de erro para a *previsão curta* são os erros das observações, os erros presentes na *estimativa a priori* utilizada durante a assimilação de dados e os erros existentes nas estimativas das estatísticas dos erros. Na prática isto significa que nas simulações é utilizado sempre o mesmo modelo para todos os membros do *ensemble*. As estatísticas estimadas por este método foram comparadas com as utilizadas operacionalmente na 4D-Var do Arpège e, em particular, com as provenientes do método de NMC. O impacto na qualidade das previsões do modelo foi também estudado (Belo Pereira e Berre, 2006).

Globalmente, verificou-se que nas estimativas do método de *ensemble* as contribuições dos fenómenos de mesoscala para os erros da *previsão curta* parecem ser acentuadas, relativamente às estimadas pelo método de NMC. Consequentemente, o comprimento de correlação (que descreve a forma como a função de autocorrelação decresce com a distância) é menor quando estimado pelo método de *ensemble*, do que quando obtido pelo método de NMC. Estes resultados são concordantes com os descritos por Fisher (2003), que considerou que estas diferenças seriam provocadas pela utilização de previsões de maior alcance no método de NMC.

No entanto, na técnica de NMC, as estimativas dos erros dependem em grande parte do incremento da análise (diferença entre a análise e a *estimativa a priori*). Além disso, verificou-se que o espectro do incremento da análise é mais largo que o espectro do erro da análise estimado pelo método de *ensemble*. Por outras palavras, as contribuições dos fenómenos de larga escala são acentuadas no espectro do incremento da análise, quando comparadas com o espectro das diferenças entre análises de membros distintos de um *ensemble*. Este resultado é consistente com as propriedades de filtragem da matriz **B**. Isto sugere que as diferenças entre as duas técnicas relativamente à representação da análise na simulação da evolução dos erros, explicam também uma parte importante das diferenças entre os dois métodos.

O estudo dos mapas de variâncias dos erros das *previsões curtas* e a comparação entre as variâncias dos erros da análise e da *previsão curta* estimados pelo método de NMC, permitiu obter resultados concordantes com os estudos de Bouttier (1994) e McNally (2000), referidos anteriormente. Por outro lado, os mapas de variâncias dos erros da *estimativa a priori* provenientes do método de *ensemble* parecem identificar algumas características interessantes, tais como os contrastes entre regiões com baixa e com elevada densidade de observações, e também as áreas de grande variabilidade atmosférica (a Zona de Convergência Inter-Tropical, por exemplo). Além disso, na técnica de *ensemble*, as diferenças entre os campos da análise de diferentes membros do *ensemble* parecem fornecer informações relevantes sobre os erros da análise, por

exemplo, relativamente à influência da densidade das observações na qualidade da análise.

As variações geográficas da função de correlação foram também examinadas. Com esta finalidade, foi desenvolvido um método económico para estimar localmente o comprimento de correlação. Aplicando este método às estimativas provenientes do método de *ensemble*, verificou-se que o comprimento de correlação é maior nos trópicos e menor nas latitudes médias (sendo este contraste máximo para as regiões ricas em observações). Este resultado é consistente com outros estudos (por exemplo, Lindzen e Fox-Rabinovitz (1989), Ingleby (2001), Bouttier (1994)). Além disso, estas estimativas derivadas do método de *ensemble* diferem das provenientes do método de NMC, principalmente para a temperatura. Em particular, os resultados sugerem que o método de NMC tende a sobrestimar o comprimento de correlação principalmente em áreas ricas em observações, tais como a América do Norte. Estes resultados sugerem que os efeitos previsíveis da densidade de observações nas estruturas das correlações dos erros da *estimativa a priori* não são bem representados pelo método de NMC.

É também importante referir que, este diagnóstico local do comprimento de correlação permite não só estudar as heterogeneidades, mas também examinar a anisotropia da função de correlação. Esta informação pode ser útil para modelar as variações espaciais (horizontais) das correlações da matriz **B**, utilizando, por exemplo, técnicas de *wavelets* (Fisher (2003), Deckmyn e Berre (2005)).

Os estudos de impacto realizados com a 4D-Var do modelo Arpège, permitiram verificar que: i) a utilização das covariâncias globais estimadas pelo método *ensemble* (quando comparadas com as provenientes da técnica de NMC) teve um impacto positivo na qualidade das previsões do modelo; ii) este impacto positivo foi ampliado, quando foram também utilizados os mapas tridimensionais das variâncias, estimadas pelo método de *ensemble* (em vez dos operacionais). Por estas razões, o método de *ensemble* passou a ser utilizado operacionalmente na 4D-Var do modelo Arpège, desde Janeiro de 2004.

Contudo, seria de esperar que a magnitude dos erros fosse subestimada pelo método de *ensemble* devido à suposição de que o modelo é perfeito. Por isso, o impacto de simular o erro do modelo na técnica de *ensemble* foi também examinado, realizando estudos de diagnóstico e de impacto na qualidade das previsões. Para este fim, admitiu-se que as incertezas associadas à discretização e às parametrizações físicas constituem importantes fontes de erros que não devem ser desprezadas.

Para simular o efeito do erro do modelo, construiu-se um conjunto de experiências, onde o modelo foi executado com resoluções diferentes e usando valores diferentes para determinados parâmetros de algumas parametrizações. As covariâncias resultantes foram comparadas com as estimadas pela técnica de *ensemble* utilizando a hipótese de que o modelo é perfeito. Os resultados sugerem que a resolução é a fonte de erro mais importante para as variáveis de pequena escala (por exemplo, vorticidade e divergência). No entanto, as incertezas associadas às parametrizações físicas são também importantes para a vorticidade e divergência nos trópicos e, em geral, para a temperatura. Este resultado é mais notório na média e baixa troposfera, nas latitudes médias e elevadas e nos níveis superiores nos trópicos.

Além disso, verificou-se que o efeito da simulação dos erros do modelo no método de *ensemble* é aumentar a variância dos erros da *previsão curta*. Para variáveis de larga escala (tais como a pressão à superfície e a temperatura), este aumento da variância está associado à amplificação das contribuições dos fenómenos de escala sinóptica para os erros da *previsão curta*. Por outro lado, para as outras variáveis (vorticidade, divergência e humidade) a simulação do erro do modelo parece acentuar as contribuições relativas dos fenómenos de mesoscala, principalmente devido à simulação do erro do modelo relacionado com a resolução. Isto tem como consequência uma redução do acoplamento entre os campos da massa e do vento.

Os resultados mostram que a referida redução do acoplamento massa/vento tem um impacto positivo na qualidade das previsões do geopotencial nos trópicos e em áreas extra-tropicais na estratosfera e na vizinhança da tropopausa.

Por outro lado, na troposfera média e alta das latitudes médias, a referida redução do acoplamento massa/vento causa uma degradação da qualidade das previsões de geopotencial. Isto indica que nestas regiões a presença na matriz **B** de um equilíbrio geostrófico relativamente forte é relevante, pois permite que a análise do geopotencial possa beneficiar das observações de vento.

Estes resultados confirmam que o acoplamento entre os campos da massa e do vento na estratosfera é mais fraco do que na troposfera, sendo também mais fraco nos trópicos do que nas latitudes médias. Assim, estes resultados sugerem também que, na especificação da matriz **B**, a dependência em função da latitude e altitude, do acoplamento entre os campos da massa e do vento, deve ser reforçada.

Quando o erro do modelo é simulado em vez de ser negligenciado, há uma melhor concordância entre os perfis verticais das estatísticas das diferenças entre as observações e a *previsão curta* e os perfis das variâncias estimadas pelo método de *ensemble*.

Finalmente, o efeito de desprezar os erros das observações no método de *ensemble* foi também estudado. Neste caso, as simulações de *ensemble* foram executadas admitindo que os erros da *estimativa a priori* são provocados principalmente pelos erros do modelo (devido às incertezas relativas à resolução e às parametrizações físicas, ou somente devido às incertezas nas parametrizações). As covariâncias resultantes destas simulações foram comparadas com as estimadas pelo método de *ensemble* aplicado nas duas outras variantes (em que se aceita que o modelo é perfeito e em que ambos os erros das observações e do modelo são simulados). Globalmente, verificou-se que a suposição de que as observações são perfeitas conduz a uma diminuição da variância do erro da *previsão curta*, principalmente devido a uma redução das contribuições dos fenómenos de larga escala. Este efeito é consistente com as propriedades de filtragem da equação da análise. Este resultado é mais evidente na alta troposfera e para a pressão à superfície.

Os resultados mostram também que a variância estimada pela técnica de *ensemble* é menor quando se supõe que as observações são perfeitas do que quando se assume que o modelo é perfeito. Há duas interpretações possíveis para este resultado. Por um lado, os erros das observações poderão ter uma contribuição maior para o erro da *previsão curta* do que as incertezas inerentes ao modelo. Por outro lado, este resultado pode também sugerir que algumas fontes importantes de erros do modelo foram subestimadas.

A dependência das estimativas do método de *ensemble* (no contexto de um modelo perfeito) em função da resolução do modelo foi também estudada. Para cumprir este objectivo construíram-se dois conjuntos de *ensembles* semelhantes, onde o mesmo modelo numérico foi utilizado, mas com resoluções diferentes para cada conjunto.

Os resultados indicam que um incremento da truncatura do modelo tem como consequência um aumento da magnitude dos erros da *previsão curta*, estimados pelo método de *ensemble*. Embora este resultado possa parecer surpreendente, é importante lembrar que à medida que a truncatura do modelo aumenta, fenómenos de menor escala podem ser simulados pelo modelo. Por outras palavras, à medida que a resolução aumenta, mais componentes da variabilidade atmosférica podem ser representadas pelo modelo. Estes resultados sugerem que a determinação da variância dos erros da *previsão curta* devido aos fenómenos de pequena escala, nos modelos globais, permanece uma questão em aberto.

Em resumo, os resultados apresentados no presente trabalho mostram que o método de *ensemble* fornece informações relevantes e úteis sobre as covariâncias dos erros da *estimativa a priori*. Além disso, a técnica de *ensemble* pode também ser aplicada a modelos de área limitada, tal como revelam os recentes estudos efectuados com o modelo ALADIN (Ştefănescu, Berre and Belo Pereira 2006, Berre, Ştefănescu and Belo Pereira 2006).

# Contents

# Introduction

It is known that the accuracy of a forecast from a Numerical Weather Prediction (NWP) model is very sensitive to errors in its initial state. One goal of data assimilation is to determine this initial state (known as analysis) in such a way that the analysis will be as close as possible to the true state of the atmosphere. Moreover, the analysis should be compatible with the model dynamics, to avoid the generation of spurious gravity waves during the forecast.

However, to obtain an accurate analysis is not straightforward. One of the main difficulties in data assimilation is caused by the fact that the degrees of freedom of the modern NWP models ($\sim 10^7$) is larger than the number of independent available observations ($\sim 10^5$). Moreover, the distribution of the observation network is not uniform in space and time. For these reasons it is not enough to perform a spatial interpolation of observations into regular grids. A prior information is needed in order to solve the undetermined problem. In other words, it is necessary to have a first guess about the atmospheric state at all grid points. In modern data assimilation schemes, this first guess (known as background) is provided by a short range (six-hour) forecast from the previous analysis.

Thus, the analysis field results from a combination of observations and a background. In this sense, data assimilation can be seen as a least squares or regression problem. The relative weights given to observations and to the background are fundamentally determined by the specified observation and background error covariances matrices. Therefore, the quality of the analysis strongly depends on the accuracy of the observation and background error covariance estimates.

The background error covariance matrix (**B** matrix) is also important because it contains the balance relationships between mass and wind fields, which allows to spread the observed information from one variable to other variables.

The oldest method used to estimate the background error statistics is known as the innovation method (e.g. Hollingsworth and Lönnberg (1986), and Lönnberg and Hollingsworth (1986)), because it is based on the covariance matrix of the innovation vector (observation minus background values). The main limitations of this method are that the data coverage is sparse over large areas of the globe (such as the oceans), and that there is no information on the scales that are smaller than the available data density (and neither on the scales that are larger than the size of the data network).

A second method used to estimate the background error statistics is based on an empirical assumption: the differences between forecasts of different lengths, but valid at the same time, have similar structures to those of the short range forecast errors. This technique is known as the NMC method, because it was first implemented in the U.S. National Meteorological Center (now renamed National Center for Environmental Prediction) (Parrish and Derber 1992). This method has the advantages of being easy to implement, and of providing some estimate for the whole globe and for all the modelled scales. This probably explains that this method was (or still is) operational in many other NWP centers: at ECMWF (Rabier *et al.* 1998), at CMC in

Canada (Gauthier *et al.* 1999), at Météo-France (Desroziers *et al.* 1995), at the UKMO (e.g. Lorenc *et al.* 2000). This method was also used in limited area models (Berre 2000; Gustafsson *et al.* 2001).

More recently, an interesting alternative technique based on an ensemble of assimilation experiments was proposed by Houtekamer *et al.* (1996). This approach, known as Ensemble Analysis method or ensemble method has been also implemented operationally at ECMWF (Fisher, 2003).

The main goal of the present work is to make the estimate of the **B** matrix more accurate, in order to improve the forecast quality of a NWP model. In order to accomplish this, the ensemble approach has been applied to the 4D-Var scheme of the global ARPEGE model.

Firstly, the ensemble method has been implemented using a perfect-model framework. The resulting covariance estimates have been compared with those that were previously operational, by performing both diagnostic and impact studies (Belo Pereira and Berre, 2006). The global (i.e. spatially averaged) ensemble covariances have been compared with those of the NMC method. The local (i.e. spatially varying) ensemble variances have also been compared with the local operational variances (the latter correspond to a static 3D map of variances, that reflects some typical data density contrasts).

Furthermore, the local ensemble correlations have been diagnosed and compared with those from the NMC method. This has been done by developing an economical method to estimate the local correlation length scales. The study of these local correlations gives interesting information about the existing heterogeneities and anisotropies.

Some emphasis has also been given to study the role of the analysis step in the simulation of the background errors. Moreover, the background and analysis errors estimated by the ensemble approach have been compared.

Secondly, the role of the model error on the background error covariances estimated by the ensemble method is investigated. It is assumed that the horizontal discretization and the uncertainties associated to the physical parametrization are significant sources of model errors. The resulting covariance estimates have been compared with those from the ensemble approach using a perfect-model assumption. Both global and local covariances have been compared.

Finally, the effect of neglecting the observation error in the ensemble method is investigated. In this case, it is assumed that the background error is caused mainly by the model error. The resulting covariances are also compared with those provided by the ensemble approach in a perfect-model context.

The structure of the thesis is the following. In chapter 1 some theoretical notions about data assimilation and the **B** matrix are outlined. The formalism of the ensemble simulation of the model state error is presented in chapter 2, where it is compared with the formalism of the NMC method and with the exact model state error evolution. The results obtained using the ensemble method with a perfect-model assumption are described in chapter 3. In chapter 4, the effect of the simulated model error in the ensemble approach is presented. In chapter 5, the background error covariance are estimated by the ensemble approach in a perfect-observation context. Finally, the main conclusions and perspectives are presented.

The full versions of three publications are also included as appendices. Belo Pereira and Berre (2006, appendix F) corresponds to a formal and experimental study of the ensemble approach applied to the ARPEGE global model. The other two publications are focused on applications to the ALADIN limited area model. Berre, Ştefănescu and Belo Pereira (2006, appendix G) deals with a comparison between the ensemble approach and two other error simulation techniques, regarding the representation of the analysis effect. Ştefănescu, Berre and Belo Pereira (2006, appendix H) concerns the evolution of the dispersion spectra and the evaluation of model differences (when applying the ensemble approach to the ALADIN model).

# Chapter 1

# Analysis Objective

## 1.1 Introduction

In order to illustrate that data assimilation is an estimation problem and that the analysis field can be seen as the optimal least-squares estimate, or as the Best Linear Unbiased Estimate (BLUE) of a true state, it is useful to consider a simple scalar example.

Assuming that one wants to estimate the temperature at a certain location and that two sources of information are available at this point: for instance, one observation and one short range model forecast (background). In this case, it is possible to estimate the temperature value from a linear combination of the two sources of information:

$$x_a = K_1 x_b + K y \tag{1.1}$$

where $K$ and $K_1$ are the weights to be estimated. Here $x_a$ is the analysis value, which represents an estimate of the true value ($\hat{x}$), $x_b$ is the background and $y$ is the observation.

It is important to mention that if the analysis is an unbiased estimate of the true state, the equation (1.1) is equivalent to the following equation:

$$x_a = x_b + K(y - x_b) \tag{1.2}$$

The analysis is an unbiased estimate of the true state if $< x_a >= \hat{x}$, where $< >$ is the time average operator. This is verified when the following conditions are satisfied:

1) the errors of the two sources of information are unbiased, which means that the expectation of background and observation errors is zero, i.e., $< e_b > = < x_b - \hat{x} > = 0$ and $< e_o > = < y - \hat{x} > = 0$.

2) the sum of the weights is equal to 1:

$$K_1 + K = 1 \tag{1.3}$$

Using the definition of analysis given by the equation (1.2), the error of the analysis (estimate) can be expressed as

$$e_a = x_a - \hat{x} = x_b - \hat{x} + K(y - \hat{x} + \hat{x} - x_b) = e_b + K(e_o - e_b) \tag{1.4}$$

The analysis error variance is defined as

$$\sigma_a^2 = \langle e_a^2 \rangle = \langle (x_a - \hat{x})^2 \rangle = \frac{1}{N} \sum_{i=0}^{N} (x_a - \hat{x})^2 \tag{1.5}$$

where $N$ is the number of analysis experiments.

Using the definition of $\sigma_a^2$ given by the previous equation and equations (1.1) and (1.3), the following expression for analysis error variance can be derived, if the observation and background errors are uncorrelated:

$$\sigma_a^2 = K_1^2\sigma_b^2 + K^2\sigma_o^2 = (1-K)^2\sigma_b^2 + K^2\sigma_o^2 \tag{1.6}$$

where $\sigma_o^2 = \langle e_o^2 \rangle$ is the observation error variance and $\sigma_b^2 = \langle e_b^2 \rangle$ is the background error variance.

If $K$ and $K_1$ are both positive, from equations (1.3) and (1.6), it can be shown that the analysis error variance satisfies the condition

$$\sigma_a^2 \leq \sigma_{max}^2 \tag{1.7}$$

where $\sigma_{max}^2$ is the largest of $\sigma_b^2$ and $\sigma_o^2$. This result is important, but the goal of the analysis is more exigent, its aim is to find the $K$ for which the analysis error variance is the lowest possible. Thus, it is necessary to derivate equation (1.6) with respect to $K$, to find the $K$ that minimizes $\sigma_a^2$:

$$\frac{\partial\sigma_a^2}{\partial K} = 2K\sigma_o^2 + 2K\sigma_b^2 - 2\sigma_b^2 = 0 \tag{1.8}$$

The second derivative $\frac{\partial^2\sigma_a^2}{\partial^2 K}$ is positive for all $K$. Thus, the value $K$ that satisfies the equation (1.8) corresponds to the value for which $\sigma_a^2$ is minimum, and is given by the following expression:

$$K = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2} \tag{1.9}$$

The weight given by the previous equation is called the optimal weight. Substituting the previous expression into equation (1.2) or into (1.1) and (1.3), the analysis field can be expressed as:

$$x_a = x_b + \frac{\sigma_b^2}{\sigma_b^2+\sigma_o^2}(y - x_b) \quad \Leftrightarrow \quad x_a = \frac{\sigma_o^2}{\sigma_b^2+\sigma_o^2}x_b + \frac{\sigma_b^2}{\sigma_b^2+\sigma_o^2}y \tag{1.10}$$

This equation shows that if the observation is very inaccurate (or absent) or if the background is very accurate ($\sigma_o^2 >> \sigma_b^2$), then the observation is ignored ($K = 0$) and the analysis is taken from the background value: $x_a = x_b$. On the contrary, if the observation is very accurate, ($\sigma_o^2 << \sigma_b^2$), then $K = 1$ and the analysis corresponds to the observed value (the maximum weight is given to the observation). If the observation and background error variances are equal, then $K = 1/2$ and $x_a = 0.5(x_b + y)$. So, in this last case, the analysis is given by a simple average of the two information sources.

If $K$ is given by the expression (1.9), then the expected analysis error becomes

$$\sigma_a^2 = \frac{\sigma_b^2\sigma_o^2}{\sigma_b^2+\sigma_o^2} \quad \Leftrightarrow \quad \frac{1}{\sigma_a^2} = \frac{1}{\sigma_o^2} + \frac{1}{\sigma_b^2} \tag{1.11}$$

From the last equation it can be seen that the accuracy of the analysis (defined as the inverse of the variance) is the sum of the accuracies of the observations and background. This shows that the analysis error variance satisfies the condition

$$\sigma_a^2 \leq \sigma_{min}^2 \tag{1.12}$$

where $\sigma_{min}^2$ is the smallest of the expected observation and background error variances. This is a much better result than (1.7). From (1.11) and (1.12) it is clear that the uncertainty about the true state is decreased due to the combination of two sources of information.

However, it is very important to remember that both expressions (1.11) and (1.12) are only true when $K$ is the optimal weight, which corresponds to the case when the error statistics are correct.

## 1.2 The analysis equation

The simple example with a single observation given in the previous section illustrates the basic concept of analysis. Nevertheless, the equation (1.2) is only valid when the observations are available at the analysis gridpoints, which generally is not the case.

The general analysis equation is

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - H(\mathbf{x}_b)) \tag{1.13}$$

where $\mathbf{x}_a = [x_a^1, x_a^2, ..., x_a^n]$ represents the analysis vector, and $\mathbf{x}_b = [x_b^1, x_b^2, ..., x_b^n]$ the background vector, where $n$ is the dimension of the model state vector. The observation vector of dimension $m$ is $\mathbf{y} = [y^1, ..., y^m]$ and $H$ is the observation operator which transforms a model state vector ($\mathbf{x}_a$ or $\mathbf{x}_b$) into the observation vector and $\mathbf{y} - H(\mathbf{x}_b) = \mathbf{d}$ is the observation increment, also called innovation. $\mathbf{K}$ is known as the gain matrix and contains the weights that control how the background field will be modified by the observation increments in order to produce the analysis field.

Under certain hypotheses (see appendix A.1) it is possible to prove that the $\mathbf{K}$ matrix which minimizes $\sigma_a^2$ is given by

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\,\mathbf{B}\,\mathbf{H}^T + \mathbf{R})^{-1} \tag{1.14}$$

where $\mathbf{R} = \langle (e_o)(e_o)^T \rangle$ and $\mathbf{B} = \langle (e_b)(e_b)^T \rangle$ are respectively the observation and the background error covariance matrices, and $\mathbf{H}$ is the linearized observation operator.

The solution of the analysis problem can be achieved by using several methods, such as the Optimal Interpolation Algorithm and the variational methods (see appendix A).

The exact analysis error covariance matrix, $\mathbf{A}$, for any $\mathbf{K}$, is:

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\,\mathbf{H})\,\mathbf{B}^*\,(\mathbf{I} - \mathbf{K}\,\mathbf{H})^T + \mathbf{K}\,\mathbf{R}^*\,\mathbf{K}^T \tag{1.15}$$

where $\mathbf{B}^*, \mathbf{R}^*$ are the exact covariance matrices.

Moreover, if $\mathbf{K}$ contains the optimal weights, it is possible to prove that the previous equation can be rewritten as:

$$\mathbf{A}^{-1} = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} \tag{1.16}$$

Thus, in this case, the analysis accuracy is the sum of the background and observation accuracies.

It can be also demonstrated that the equation (1.16) is equivalent to:

$$\mathbf{A} = \mathbf{B} - \mathbf{B}\mathbf{H}^T(\mathbf{H}\,\mathbf{B}\,\mathbf{H}^T + \mathbf{R})^{-1}\mathbf{H}\,\mathbf{B} \tag{1.17}$$

Since the autocovariance matrices $\mathbf{B}$ and $\mathbf{R}$ are both real positive definite, then the matrix $\mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}\mathbf{H}\mathbf{B}$ is also real positive definite. Therefore, $\mathbf{A} \leq \mathbf{B}$ (or $Tr(\mathbf{A}) \leq Tr(\mathbf{B})$, where $Tr$ is the trace operator), i.e., the uncertainty of the analysis is smaller than the background uncertainty.

It is interesting to notice that equations (1.14) and (1.16) are vector generalizations of the scalar expressions (1.9) and (1.11), respectively.

## 1.3 Background Error Covariance Matrix

The background error covariance is defined as:

$$\mathbf{B} = < (e_b)(e_b)^T > = < (\mathbf{x}_b - \hat{\mathbf{x}})(\mathbf{x}_b - \hat{\mathbf{x}})^T > \tag{1.18}$$

where $\hat{\mathbf{x}}$ represents the true state of the field, $\mathbf{x}_b$ the background field and $< >$ the time average operator.

In the univariate case, the diagonal elements of $\mathbf{B}$ correspond to the error variances $(\sigma_b^i)^2$ at the different gridpoints $i$. The off-diagonal elements of $\mathbf{B}$ correspond to the covariances between

the background errors at different gridpoints $i$ and $j$: $B_{ij} = C^b(i,j) = \sigma_b^i \sigma_b^j \rho_{ij}$, where $\rho_{ij}$ is the correlation between the background errors at different gridpoints $i$ and $j$. Furthermore, the correlation matrix is symmetric, that is, $\rho_{ij} = \rho_{ji}$.

In the multivariate case, the matrix **B** contains also the cross-covariances between different parameters, such as vorticity, divergence, temperature and surface pressure. In data assimilation, this allows the mass observations to influence the wind analysis, and vice-versa. The operational multivariate formulation of the matrix **B** is described in appendix B. The remainder of the text will focus on the univariate case for sake of simplicity.

### 1.3.1 The importance of the Background Error Covariance Matrix

As said before, the analysis problem at a given time is underdetermined because the dimension of modern NWP models is larger than the number of independent available observations. Moreover, there is often a lack of observations over oceans. So, the background is used as a preliminary estimate of the true atmospheric state. Furthermore, the **B** matrix a is key element in determining the way to correct the background with observations. Therefore, in data assimilation, it is important to have a good estimate of the background error statistics.

#### A) Example with two stations

It is supposed that it is necessary to estimate the temperature at Lisbon and Porto and that both observations and background are available in the two stations. It is assumed that the instrument type is the same in both stations and so the error variance is $\sigma_o^2$ for the two stations. It is also considered that the observation errors are uncorrelated. The background error variance is $(\sigma_b^L)^2$ at Lisbon and $(\sigma_b^P)^2$ at Porto. The correlation of background error between Porto and Lisbon is given by $\rho$. In this case, where $m = 2$, $n = 2$ and $H = I$, the analysis equation (1.13) becomes

$$\begin{pmatrix} x_a^L \\ x_a^P \end{pmatrix} = \begin{pmatrix} x_b^L \\ x_b^P \end{pmatrix} + \begin{pmatrix} K_{LL} & K_{LP} \\ K_{PL} & K_{PP} \end{pmatrix} \begin{pmatrix} y^L - x_b^L \\ y^P - x_b^P \end{pmatrix} \tag{1.19}$$

which is equivalent to the two equations for the analyses at Lisbon and Porto:

$$x_a^L = x_b^L + K_{LL}(y^L - x_b^L) + K_{LP}(y^P - x_b^P) \tag{1.20}$$

$$x_a^P = x_b^P + K_{PL}(y^L - x_b^L) + K_{PP}(y^P - x_b^P) \tag{1.21}$$

and according to equation (1.14), the Gain matrix in this case is

$$\mathbf{K} = \mathbf{B}(\mathbf{B} + \mathbf{R})^{-1} \tag{1.22}$$

where

$$\mathbf{R} = \begin{pmatrix} \sigma_o^2 & 0 \\ 0 & \sigma_o^2 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} (\sigma_b^L)^2 & \rho\sigma_b^L\sigma_b^P \\ \rho\sigma_b^P\sigma_b^L & (\sigma_b^P)^2 \end{pmatrix}. \tag{1.23}$$

Combining the previous expressions with the equation (1.22), the gain matrix can be written as:

$$\mathbf{K} = \frac{1}{(1+\alpha_P)(1+\alpha_L) + (\rho)^2} \begin{pmatrix} (1+\alpha_P) - (\rho)^2 & \rho\sqrt{\alpha_P}\sqrt{\alpha_L} \\ \rho\sqrt{\alpha_P}\sqrt{\alpha_L} & (1+\alpha_L) - (\rho)^2 \end{pmatrix} \tag{1.24}$$

where $\alpha_P = \sigma_o^2/(\sigma_b^P)^2$ and $\alpha_L = \sigma_o^2/(\sigma_b^L)^2$.

If the background errors at Lisbon and at Porto are uncorrelated ($\rho = 0$), then the Gain matrix becomes

$$\mathbf{K} = \begin{pmatrix} (1+\alpha_L)^{-1} & 0 \\ 0 & (1+\alpha_P)^{-1} \end{pmatrix} \tag{1.25}$$

The last equation is equivalent to

$$K_{LL} = \frac{(\sigma_b^L)^2}{(\sigma_b^L)^2 + \sigma_o^2}, \quad K_{PP} = \frac{(\sigma_b^P)^2}{(\sigma_b^P)^2 + \sigma_o^2}, \quad K_{PL} = K_{LP} = 0 \tag{1.26}$$

Thus, substituting the last equation into equations (1.20) and (1.21), it can be seen that the analysis at Lisbon is not influenced by observations at Porto. It is only influenced by observations at Lisbon, and the analysis at Porto depends only on observations at Porto. Moreover, it is also clear that observations and background in the two stations will be combined according to their relative accuracies, as in the scalar case (see section 1.1).

When the correlation $\rho$ increases from zero, the terms $K_{PL}$ and $K_{LP}$ increase in magnitude. Therefore, the analysis at Lisbon is more influenced by an observation at Porto and also the weight of an observation at Lisbon for the analysis at Porto increases. The maximum weight that can be given to an observation at Porto for an analysis at Lisbon is when $\rho = 1$ and

$$K_{LP}^{max} = \frac{\sqrt{\alpha_P}\sqrt{\alpha_L}}{(1+\alpha_P)(1+\alpha_L)+1}$$

and in this case, the observation at Lisbon receives a weight of

$$K_{LL} = \frac{\alpha_P}{(1+\alpha_P)(1+\alpha_L)+1}$$

for the analysis at Lisbon and $K_{LP}^{max}/K_{LL} = \sqrt{\alpha_L}/\sqrt{\alpha_P}$. Since the observation error variances were assumed to be the same at the two stations, if the background error variance is also the same at both stations, then $\alpha_P = \alpha_L$ and the observations at Porto and Lisbon have the same weight for the analysis at Lisbon. If the forecast at Porto is very good relatively to the observations and better than at Lisbon, i.e, $(\sigma_b^P)^2 < (\sigma_b^L)^2 \simeq \sigma_o^2$, then $K_{LP}^{max} < K_{LL}$, which means that one observation at Porto has less weight for the analysis at Lisbon than one observation at Lisbon. On the contrary, if the forecast at Lisbon is quite good compared with observations and with the forecast at Porto, i.e., $(\sigma_b^L)^2 < (\sigma_b^P)^2 \simeq \sigma_o^2$, then $K_{LL} < K_{LP}^{max}$ and therefore one observation at Porto has more weight for the analysis at Lisbon than one observation at Lisbon.

So, it is obvious that observations at one location can influence the analysis at other locations. Moreover, this influence is much determined by the standard deviation $\sigma_b$ and by the background error correlation, $\rho$, which is the correlation of background errors between the observation location and the analysis location.

### B) Analysis at one station with an observation in another station

A single observation is supposed to be available, for instance at Lisbon. In addition, the background error variance is assumed to be the same at Lisbon and Porto:

$$\langle (e_b^L)^2 \rangle = \langle (e_b^P)^2 \rangle = \sigma_b^2 \tag{1.27}$$

The correlation between the background error at Lisbon and Porto is:

$$\rho = \frac{\langle (e_b^P)(e_b^L) \rangle}{\sigma_b^P \sigma_b^L} \tag{1.28}$$

It is also assumed that observation and background errors are uncorrelated:

$$\langle (e_b^L)(e_o^L) \rangle = \langle (e_b^P)(e_o^L) \rangle = 0 \tag{1.29}$$

For this case, the analysis at Porto can be determined by

$$x_a^P = x_b^P + K_{PL}(y^L - x_b^L) \tag{1.30}$$

In this case, the $K_{PL}$ which minimizes the analysis variance is:

$$K_{PL} = \rho \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2} = \frac{\rho}{1+\alpha} \tag{1.31}$$

where $\alpha = \sigma_o^2/\sigma_b^2$.

Thus, this confirms that the weight given to an observation for the analysis in another location does not depend only on the ratio between the observation and background accuracies. It depends also on the distance between the observation and the analysis grid point. For instance, if this distance is so large that the correlation is 0, then the weight is 0 and the analysis error variance will be the background error variance. Finally, it also depends on the way the background error correlation varies with distance.

### C) The effect of misspecification of the background error covariance

It is also relevant to consider the case when the specified weight (in a single observation case) is:

$$\widetilde{K} = \frac{\widetilde{\sigma}_b^2}{\widetilde{\sigma}_b^2 + \sigma_o^2} \tag{1.32}$$

where $\widetilde{\sigma}_b^2$ is an incorrectly specified background error variance, while the correct error variance is $\sigma_b^2$.

Then the analysis equation (1.10) becomes

$$x_a = \frac{\sigma_o^2}{\widetilde{\sigma}_b^2 + \sigma_o^2} x_b + \frac{\widetilde{\sigma}_b^2}{\widetilde{\sigma}_b^2 + \sigma_o^2} y \tag{1.33}$$

and according to equation (1.11), the perceived expected analysis error variance is

$$\widetilde{\sigma}_a^2 = \frac{\widetilde{\sigma}_b^2 \sigma_o^2}{\widetilde{\sigma}_b^2 + \sigma_o^2} \tag{1.34}$$

However, since $\widetilde{K}$ is not the optimal weight (because $\widetilde{\sigma}_b^2$ is not the real background error variance), then the real analysis error variance is given by the equation (1.6), which in this case can be rewritten as:

$$\sigma_a^2 = \left(\frac{\sigma_o^2}{\widetilde{\sigma}_b^2 + \sigma_o^2}\right)^2 \sigma_b^2 + \left(\frac{\widetilde{\sigma}_b^2}{\widetilde{\sigma}_b^2 + \sigma_o^2}\right)^2 \sigma_o^2 = \frac{\widetilde{\sigma}_b^4 \sigma_o^2 + \sigma_o^4 \sigma_b^2}{(\sigma_o^2 + \widetilde{\sigma}_b^2)^2} \tag{1.35}$$

According to Daley (1991), it is interesting to consider two extreme cases:

In the first case, $\sigma_b^2 \ll \sigma_o^2 \ll \widetilde{\sigma}_b^2$. Thus, the analysis would give more weight to the observations, and from expressions (1.34) and (1.35) it can be seen that the perceived and real analysis error variances are $\sigma_a^2 \approx \widetilde{\sigma}_a^2 \lesssim \sigma_o^2$. So, the expected analysis error variance is only smaller than the larger of $\sigma_b^2$ and $\sigma_o^2$, because in the analysis a very good background (comparing with observations) was ignored. If $\sigma_b^2$ had been correctly specified, the observation would have been discarded and the analysis error variance would have been reduced to its smallest possible value.

In the second case, $\widetilde{\sigma}_b^2 \ll \sigma_o^2 \ll \sigma_b^2$. Then, in this case, any analysis scheme would give more weight to the background, which is very inaccurate. Moreover, as it is assumed that $\widetilde{\sigma}_b^2 \ll \sigma_o^2$, which implies that $\sigma_o^2 \lesssim \sigma_o^2 + \widetilde{\sigma}_b^2$, then it is clear from equation (1.34) that $\widetilde{\sigma}_a^2 \lesssim \widetilde{\sigma}_b^2$. So, the perceived analysis error variance would be impressively small. However, $\sigma_a^2$ would be unacceptably large, because $\sigma_a^2 \lesssim \sigma_b^2$. This type of error is much more serious than the one done in the first case, because in this case the background field is very inaccurate, while in the first case the observations are less accurate than the background, but are not extremely inaccurate ($\sigma_o^2 \ll \widetilde{\sigma}_b^2$). Thus, when a background is used in the objective analysis, it is wiser to overestimate the expected background error variance than to underestimate it.

### 1.3.2 Length scale of autocorrelation

In the current NWP models, the model state dimension is $n \sim 10^7$. Thus, the dimension of $\mathbf{B}$ matrix is $\sim 10^7 \times 10^7$, which is too big to be computed. For this reason, it is necessary to make some assumptions, which simplify the computation of this matrix and reduce its computational cost.

The background error covariances of a 3-dimensional field in the grid point space of a model can be simplified if the vertical and horizontal structures are separable. This means that

$$\mathbf{B}([x_i, y_i, z_i], [x_j, y_j, z_j]) = \sigma_b(x_i, y_i, z_i)\, \sigma_b(x_j, y_j, z_j)\, C_H^b([x_i, y_i], [x_j, y_j])\, C_V^b(z_i, z_j) \tag{1.36}$$

where $\sigma_b$ is the standard deviation of the background error, $C_H^b$ is the horizontal correlation, which is independent of height and $C_V^b$ is the vertical correlation which is independent of the horizontal coordinates.

Generally, the horizontal background error correlations are assumed to be homogeneous. This means that the correlation between two points in physical space depends only on the horizontal relative displacement, $\tilde{\mathbf{r}} = (x_j - x_i, y_j - y_i)$, and not on the absolute locations $(x_i, y_i)$ and $(x_j, y_j)$. So, under homogeneous conditions, the correlation can be written as a function of only two parameters: the norm of the horizontal relative displacement ($r$) and the direction (given by the angle, $\theta$):

$$C_H^b([x_i, y_i], [x_j, y_j]) \approx C^b(r, \theta) \tag{1.37}$$

where $r^2 = (x_j - x_i)^2 + (y_j - y_i)^2$ and $\theta = tan^{-1}\big((y_j - y_i)/(x_j - x_i)\big)$.

If the background error correlation does not depend either on the direction, i.e., it depends only on the horizontal distance $r$ (between the two points), it is said to be isotropic:

$$C_H^b([x_i, y_i], [x_j, y_j]) \approx C^b(r) \tag{1.38}$$

otherwise, it is anisotropic.

Several studies show that the standard deviation (square root of the variance) of background errors is not constant over the globe. Therefore, the horizontal covariance for geopotential background error is not homogeneous or isotropic. A less restrictive and more reasonable assumption is to assume that only the correlations are homogeneous or isotropic, because it gives the possibility of introducing spatial variation of the background error standard deviation.

The correlation matrix is defined by:

$$\rho = \frac{\mathbf{B}([x_i, y_i, z_i], [x_j, y_j, z_j])}{\sigma_b(x_i, y_i, z_i)\sigma_b(x_j, y_j, z_j)} \tag{1.39}$$

or

$$\mathbf{B} = \mathbf{D}\,\rho\,\mathbf{D} \tag{1.40}$$

where $\mathbf{D}$ is the diagonal matrix of standard deviations. The standard deviations are estimated for each point on the grid and for each variable.

Daley (1991) defined the correlation length scale ($L$) as a measure of the inverse local curvature of the autocorrelation $\rho$. In the one-dimensional (1D) homogeneous case, it is defined as:

$$L^2 = -\left.\frac{\rho(r)}{d^2\rho/dr^2}\right|_{r=0} \tag{1.41}$$

If $\frac{d\rho}{dr}$ is continuous at $r = 0$, it must be zero, because $\rho$ decreases when $r$ increases from zero. Since $L^2$ is positive, the second derivative of $\rho(r)$ must be negative at $r = 0$. Moreover, a 1D homogeneous autocorrelation function can be shown to be also isotropic. Therefore, $\rho$ is symmetric about $r = 0$, and the first derivative should be zero at $r = 0$. Thus, the autocorrelation function is approximately parabolic near $r = 0$. In the two-dimensional (2D) case, $L$ is defined by

$$L^2 = -\frac{2\rho}{\nabla^2\rho}\bigg|_{r=0} \tag{1.42}$$

where $\nabla^2$ is the Laplacian operator.

$L$ describes the sharpness of $\rho$ (i.e. how it decays with distance from its initial value 1). For a sharp autocorrelation function, the curvature is high and the length scale must be small. Similarly, if the curvature is low, the autocorrelation function is wide and the length scale is large.

In practice, the correlation length scale gives an idea about the distance over which the influence of the observation extends. Consequently, the resolution of the analysis increments is greatly determined by the horizontal correlation lengths used in the background error covariance model.

It is interesting to notice that $L$ reflects mostly short-distance correlations and it is very sensitive to the small scale part of the spectrum. This explains why it is sometimes called microscale (Daley, 1991).

### 1.3.3 Covariance spectra

As said before, under the homogeneity and isotropy assumptions, the background error covariance between two points is a function of only their distance over the sphere:

$$< e_b(\lambda,\mu)e_b(\lambda',\mu') >= f(r)$$

where $\mu$ represents the sinus of latitude and $\lambda$ the longitude. Hereafter, it is assumed that $e_b$ is unbiased.

Besides, it is possible to express $f(r)$ in terms of spherical harmonics:

$$f(r) = \sum_{n=0}^{N}\sum_{m=-n}^{m} a_n^m P_n^m(cos\gamma)e^{im\lambda} \tag{1.43}$$

where $\gamma$ is the angular separation between the two points and $P_n^m(cos\gamma)$ is the Legendre polynomial of first kind with degree $m$ (and $m$ is the zonal wave number) and order $n$ (and $n$ is the total wave number).

Moreover, under the homogeneity and isotropy assumptions, the covariance function can be written as

$$f(r) = \sum_{n=0}^{N} a_n P_n^0(cos\gamma) \tag{1.44}$$

and the background variance on the sphere is given by

$$f(r=0) = \sum_{n=0}^{N} a_n P_n^0(1) = \sum_{n=0}^{N} a_n\sqrt{(2n+1)} \tag{1.45}$$

The last equality in the previous equation results from applying the normalization to the Legendre polynomials, which leads to $P_n^0(1) = \sqrt{(2n+1)}$.

It was demonstrated (Boer, 1983) that, under the previous assumptions, the background error covariance can be written as:

$$< (e_b)_n^m[(e_b)_{n'}^{m'}]^* >= \delta_n^{n'}\delta_m^{m'}B_n = \begin{cases} B_n = \frac{a_n}{\sqrt{(2n+1)}} & \text{if} \quad (n,m) = (n',m') \\ 0 & \text{if} \quad (n,m) \neq (n',m') \end{cases} \tag{1.46}$$

and $B_n$ is known as the modal variance, which is a function of total wave number only, i.e., it is independent of the zonal wave number.

The meaning of the last equation is that under the homogeneity and isotropy assumptions, the **B** matrix is diagonal in the spectral space. This is a very important result because this means that the calculation of a full matrix of two-dimensional (i.e, horizontal) covariances in grid point space is equivalent to the computation of a diagonal matrix (containing spectral variances) in spectral space. Therefore, when the previous assumptions are used, it is more efficient to calculate covariances in spectral space than in gridpoint space.

Combining the equations (1.45) and (1.46), it is possible to write the global variance of the background error in the following way:

$$\sigma_b^2 =< e_b^2 >= \sum_{n=0}^{N} \sum_{m=-n}^{m} \langle [(e_b)_n^m]^2 \rangle = \sum_{n=0}^{N} B_n(2n+1) = \sum_{n=0}^{N} P_n \tag{1.47}$$

where $P_n = a_n\sqrt{(2n+1)}$ is the total variance for a given total wavenumber $n$ and provides the covariance power spectrum. The correlation spectra ($h_n$) are deduced from the covariance spectra by normalizing them by the global variance: $h_n = P_n/\sigma_b^2$.

The last equation establishes the correspondence between the background error variance in gridpoint space and the covariance power spectrum.

In the last years, several studies with global models (Courtier *et al.*, 1998; Rabier *et al.*, 1998) have shown that it is easy to formulate non-separable structure functions in spectral space.

The total vertical covariance between two levels $l$ and $k$ can be written as:

$$B^{kl} = \frac{1}{4\pi} \int_0^{2\pi} \int_{-1}^{1} e_b^l(\lambda,\mu) e_b^{k*}(\lambda,\mu) d\lambda d\mu = \sum_{n=0}^{N} \sum_{m=-n}^{m} \langle (e_b^k)_n^m [(e_b^l)_n^m]^* \rangle = \sum_{n=0}^{N} B_n^{kl}$$

where

$$B_n^{kl} = \sum_{m=-n}^{m} \langle (e_b^k)_n^m [(e_b^l)_n^m]^* \rangle$$

is the vertical covariance for the wavenumber $n$. The total covariance can also be written as

$$B^{kl} = \sum_{n=0}^{N} \sqrt{B_n^{kk}} \sqrt{B_n^{ll}} \quad V_n^{kl}$$

where

$$V_n^{kl} = \frac{B_n^{kl}}{\sqrt{B_n^{kk}}\sqrt{B_n^{ll}}}$$

is the vertical correlation between levels $k$ and $l$ for the wavenumber $n$ and $\sqrt{B_n^{kk}}$ is the standard deviation for the wavenumber $n$ at level $k$.

Finally, using the definition of length scale (eq. 1.42), from Daley (1991), $L$ can also efficiently be computed in spectral space, according to:

$$L^2 = -2 \frac{\sum_{n=0}^{N} B_n}{\sum_{n=0}^{N} n(n+1) B_n} \tag{1.48}$$

### 1.3.4 Filtering properties

Following Hollingsworth (1987) and Daley (1991) it is possible to illustrate the filtering properties of an analysis algorithm.

Assuming that the analysis gridpoints and observation stations coincides ($H = I$), the analysis equation (1.13) becomes:

$$\mathbf{x}_a - \mathbf{x}_b = \mathbf{K}(\mathbf{y} - \mathbf{x}_b) \tag{1.49}$$

where

$$\mathbf{K} = \mathbf{B}(\mathbf{B}+\mathbf{R})^{-1} = (\mathbf{I}+\mathbf{R}\mathbf{B}^{-1})^{-1} \tag{1.50}$$

In order to illustrate the filtering properties of (1.49) it is necessary to examine the eigenstructures of the covariance matrices involved in $\mathbf{K}$.

It is possible to show that if $e$ is an eigenvector of the matrix $\mathbf{R}\mathbf{B}^{-1}$ with an associated eigenvalue $\nu$, then $\nu$ must be real and positive, because $\mathbf{R}$ and $\mathbf{B}$ are both symmetric positive definite matrices (see Daley 1991, section 4.5). In addition, in this case $e$ is also an eigenvector of $\mathbf{K}$ and

$$\mathbf{K}e = (\mathbf{I}+\mathbf{R}\mathbf{B}^{-1})^{-1}e = (1+\nu)^{-1}e \tag{1.51}$$

Consequently, because $\nu$ is positive, $0 < (1+\nu)^{-1} < 1$ and all the eigenvalues of $\mathbf{K}$ lie between zero and one.

It is interesting to consider the special case in which the background and observation error matrices commute, then $\mathbf{B}\mathbf{R} = \mathbf{R}\mathbf{B}$. Moreover, supposing that $e$ is an eigenvector of $\mathbf{B}$ and $\mathbf{R}$ with associated eigenvalues $\mu_B^2$ for $\mathbf{B}$ and $\mu_o^2$ for $\mathbf{R}$:

$$\mathbf{B}e = \mu_B^2 e \quad \text{and} \quad \mathbf{R}e = \mu_o^2 e \tag{1.52}$$

Then, equation (1.51) can be written as

$$\mathbf{K}e = \left[1 + \frac{\mu_o^2}{\mu_B^2}\right]^{-1} e \tag{1.53}$$

It is important to mention that $\mu_B^2$ and $\mu_o^2$ are the expected background and observation error variances for the eigenvector $e$ (see Appendix D, Daley 1991).

If observation errors are spatially uncorrelated (with expected variance $\sigma_o^2$, independent of the observation station), then $\mathbf{R} = \sigma_o^2\mathbf{I}$ and consequently $\mu_o^2 = \sigma_o^2$.

Suppose that $\mathbf{B} = \sigma_b^2\rho$, where $\rho$ is the background error correlation matrix. Furthermore, assume that $e$ is an eigenvector of $\rho$ with the corresponding eigenvalue $\lambda$, i.e., $\rho e = \lambda e$, then combining the last expressions with equations (1.52) and (1.53) it is possible to write

$$\mu_B^2 = \sigma_b^2\lambda \quad \text{and} \quad \mathbf{K}e = \left[1 + \frac{\varepsilon_o^2}{\lambda}\right]^{-1} e \tag{1.54}$$

where $\varepsilon_o^2 = \sigma_o^2/\sigma_b^2$.

If the observation increment can be written as a superposition of eigenvectors of $\rho$:

$$\mathbf{d} = \mathbf{y} - \mathbf{x}_b = \sum_{i=1}^{N} c_i e_i \tag{1.55}$$

then

$$\mathbf{x}_a - \mathbf{x}_b = \mathbf{K}\mathbf{d} = \sum_{i=1}^{N} c_i \mathbf{K} e_i = \sum_{i=1}^{N} c_i \left[1 + \frac{\varepsilon_o^2}{\lambda}\right]^{-1} e_i \tag{1.56}$$

Because $[1 + \frac{\varepsilon_o^2}{\lambda}]^{-1}$ varies between zero and one, according to equation (1.56) the analysis algorithm always behaves as a filter of the observation increments.

Equation (1.56) also indicates that the eigenvectors of $\rho$ with the largest eigenvalues ($\lambda >> \varepsilon_o^2$) are filtered the least, and those with the smallest eigenvalues ($\lambda << \varepsilon_o^2$) are filtered the most. In addition, when $\lambda >> \varepsilon_o^2$ for all eigenvectors of the correlation matrix $\rho$, the observations are well fitted and there is little filtering. On the other hand, when $\lambda << \varepsilon_o^2$ for all the eigenvectors of $\rho$, the observations are ignored and the analysis is similar to the background field.

Moreover, it can be demonstrated that the eigenvectors of $\rho$ have decreasing eigenvalues as their effective scale decreases (Daley, 1991, p. 128). Consequently, for uncorrelated observation

errors, $\mathbf{K}$ acts like a low-pass filter, i.e., it tends to filter the smaller scales (of the observation increment) most strongly.

It is also important to remember that the analysis equation can be also written as:

$$\mathbf{x}_a = \mathbf{K}\mathbf{y} + (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{x}_b$$

where contrarily to $\mathbf{K}$, $\mathbf{I} - \mathbf{K}\mathbf{H}$ acts like a high pass filter.

Thus, the larger scales of the background field and the small scales of observations are filtered the most. In other words, the analysis can be seen as the combination of the large scale part of the observation field $\mathbf{y}$ and of the small scale part of the background field $\mathbf{x}_b$.

The equation (1.56) is valid when the observation errors are spatially uncorrelated. When the observation errors are spatially correlated (e.g. for satellite data), Daley (1991) has shown that the scale selectivity of the analysis filter depends on the ratio between the correlation length scales of the observation error and of the background error. If the correlation length scale of the observation error ($L_o$) is smaller than $L$, the analysis algorithm behaves like a low-pass filter on the observation increments, like for the uncorrelated observation errors. However, when $L_o = L$, the filter $\mathbf{K}$ is independent of the scale.

On the other hand, when $L_o > L$, $\mathbf{K}$ acts like a high-pass filter on the observation increments. Therefore, the analysis uses mostly the information from observations at small scales and from the background at large scales. This can be explained because when $L_o > L$, the observation error is mostly in large scales and only the smaller scales of the observations are reliable.

# Chapter 2

# Simulation methods used for the estimation of background error covariances

*This chapter corresponds essentially to the formal part of Belo Pereira and Berre (2006). It has been extended by including some additional remarks about the model error representation.*

## 2.1 Introduction

Most of the major operational NWP centers use variational methods to combine observations with a background state provided by a short range forecast, in order to produce an optimal estimate of the atmospheric state. However, this estimate is only optimal if both observation and background error covariance matrices are correctly specified in the analysis. Therefore, an accurate specification of these covariance matrices is very important for the quality of the assimilation system. For instance, the accuracy of the background error covariance is important for the assimilation of satellite sounding data such as from the TIROS[1] Operational Vertical Sounder (TOVS) (Smith *et al.*, 1979, Eyre, 1989b). The poor vertical resolution of these data implies that the analysis strongly depends on the vertical correlation of background temperature (and humidity) error to distribute the radiance information in the vertical (McNally, 2000). Nevertheless, the estimation of the background error statistics is not straightforward, since the truth is never exactly known.

This chapter presents the formal comparison between two approaches which can be used to estimate the background error covariances in the NWP operational context.

One method is known as the NMC method, which relies on the assumption that the background error can be estimated by computing the differences between forecasts of different ranges, but valid at the same time.

The other technique is called ensemble method, which was firstly proposed by Houtekamer *et al.* (1996). It is based on an ensemble of assimilation experiments: for each ensemble member, the observations can be randomly perturbed, as well as the physical parametrizations.

The formalism of the method based on an ensemble of assimilation experiments is briefly described. It will be also compared with the formalism of the NMC method. A linear framework

[1]Television Infra-Red Observation Satellite

will be considered to make the derivations simpler. The reader is referred to Berre *et al.* (2006) for more detailed derivations (note also that a formalism in a non linear framework is described in Žagar *et al.* (2004)). In the current chapter, the main features that are important for the interpretation of the experimental results will be summarized. This concerns for instance the differences between the two methods regarding the analysis step representation, as it will be seen.

Firstly, the exact evolution of the model state error is presented. The way this error evolution is simulated in the ensemble approach is then explained and it is finally compared with the corresponding formalism of the NMC method. The true errors will be noted $e$, while the simulated errors will be noted $\varepsilon$ and $\tilde{\epsilon}$ respectively in the ensemble and NMC methods. The experimental environment of the ensemble simulations is described in the next chapter.

## 2.2 The exact model state error evolution

An assimilation cycle involves a succession of analyses and forecasts. This process implies a specific evolution for the model state errors.

Starting from an uncertain analysis $\mathbf{x}_a^i$ (valid at time $t_i$), one will obtain an uncertain six-hour (6h) forecast $\mathbf{x}_b^{i+1}$ (valid at time $t_{i+1} = t_i + 6h$) by integrating the NWP model: $\mathbf{x}_b^{i+1} = \mathbf{M}\,\mathbf{x}_a^i$, where $\mathbf{M}$ is the operator that corresponds to the 6h integration provided by the forecast model. If the analysis and the forecast model were perfect, the last equation would become $\hat{\mathbf{x}}^{i+1} = \hat{\mathbf{M}}\,\hat{\mathbf{x}}^i$, where $\hat{\mathbf{M}}$ is the exact 6h forecast operator (which represents a perfect forecast model) and $\hat{\mathbf{x}}$ is the true state. Thus, the background error can be expressed as:

$$e_b^{i+1} = \mathbf{x}_b^{i+1} - \hat{\mathbf{x}}^{i+1} = \mathbf{M}\,\mathbf{x}_a^i - \mathbf{M}\,\hat{\mathbf{x}}^i + \mathbf{M}\,\hat{\mathbf{x}}^i - \hat{\mathbf{M}}\,\hat{\mathbf{x}}^i = \mathbf{M}\,(\mathbf{x}_a^i - \hat{\mathbf{x}}^i) + (\mathbf{M} - \hat{\mathbf{M}})\,\hat{\mathbf{x}}^i$$

Consequently, the evolution of the analysis errors into the background errors, $e_b^{i+1}$, corresponds to the following equation:

$$e_b^{i+1} = \mathbf{M}e_a^i + e_m^{i+1} \tag{2.1}$$

where $e_a^i = \mathbf{x}_a^i - \hat{\mathbf{x}}^i$ is the analysis error and $e_m^{i+1}$ is the accumulated model error (during the 6h period). The model error is defined as:

$$e_m^{i+1} = \mathbf{M}\,\hat{\mathbf{x}}^i - \hat{\mathbf{x}}^{i+1} = \mathbf{M}\,\hat{\mathbf{x}}^i - \hat{\mathbf{M}}\,\hat{\mathbf{x}}^i = (\mathbf{M} - \hat{\mathbf{M}})\,\hat{\mathbf{x}}^i$$

The 6h forecast field $\mathbf{x}_b^{i+1}$ will then be used as the background for a new analysis (that is valid at time $t_{i+1}$, according to the equation (1.13), which will use also the observation vector $\mathbf{y}^{i+1}$). As the analysis equation also applies to the states $\hat{\mathbf{x}}$ and $\mathbf{H}(\hat{\mathbf{x}})$, it can be shown that, the evolution of the background errors $e_b^{i+1}$ into the analysis errors $e_a^{i+1}$ corresponds to the similar following equation:

$$e_a^{i+1} = e_b^{i+1} + \mathbf{K}(e_o^{i+1} - \mathbf{H}e_b^{i+1}) = \mathbf{K}e_o^{i+1} + (\mathbf{I} - \mathbf{KH})e_b^{i+1} \tag{2.2}$$

where $\mathbf{K}$ is the classical gain matrix ($\mathbf{K} = \mathbf{BH}^T(\mathbf{HBH}^T + \mathbf{R})^{-1}$, where $\mathbf{B}, \mathbf{R}$ are respectively the specified background and observation error covariance matrices) and $\mathbf{H}$ is the observation operator. This operator transforms the model variables into observed variables and performs the spatial interpolation from the model grid to the observation locations.

The observation vector can be written as $\mathbf{y} = \hat{\mathbf{y}} + e_{me}$, where $\hat{\mathbf{y}}$ is the exact atmospheric state in the observation space, and $e_{me}$ is the measurement error. It is also possible to express $\mathbf{y}$ in the following way:

$$\mathbf{y} = \mathbf{H}(\hat{\mathbf{x}}) + \hat{\mathbf{y}} - \mathbf{H}(\hat{\mathbf{x}}) + e_{me} = \mathbf{H}(\hat{\mathbf{x}}) + e_r + e_{me} = \mathbf{H}(\hat{\mathbf{x}}) + e_o$$

where $e_r$ is the error of representativeness, which reflects the fact that the operator $\mathbf{H}$ is not perfect. The observation error, $e_o^{i+1}$, is the sum of the measurement error and of the representativeness error.

It is interesting to note that the analysis equation, which is applied to observation and background fields is the same as the equation applied to the corresponding errors (see equations 1.13 and 2.2). Furthermore, since $\mathbf{K}$ can be seen as a low pass filter, and $\mathbf{I}-\mathbf{KH}$ acts like a high pass filter, the analysis error can be seen as the combination of a large scale part of the observation error $e_o$ and of a small scale part of the background error $e_b$ (see section 1.3.4).

## 2.3 The ensemble simulation of the error evolution

The ensemble method may be seen as a technique that intends to simulate the exact error evolution, by applying an intermittent data assimilation cycle to an ensemble of perturbed experiments. This approach is illustrated schematically in figure 2.1. For each member $n$ (for $n = 1, ..., N$), the observations are perturbed by adding independent random Gaussian perturbations $\delta_o(n)$, which are drawn from the specified observation error covariance matrix $\mathbf{R}$. At time $t_{i+1} = t_i + 6h$, the analysis for each member is

$$\mathbf{x}_a^{i+1}(n) = \mathbf{x}_b^{i+1}(n) + \mathbf{K}(\mathbf{y}^{i+1}(n) - \mathbf{H}\mathbf{x}_b^{i+1}(n)) \tag{2.3}$$

where $\mathbf{y}^{i+1}(n) = \mathbf{y}^{i+1} + \delta_o^{i+1}(n)$ are the perturbed observations and $\mathbf{x}_b^{i+1}(n)$ represents the background field for each member $n$.

Thus, at a given time an ensemble of (perturbed) analyses is available. Moreover, the dispersion of this ensemble of analyses, given by the analysis differences $\varepsilon_a$, can be seen as an estimate of the analysis error, $e_a^i$, due to the uncertainties in the observation and background fields.

From equation (2.3) it can be shown that the background differences $\varepsilon_b^{i+1} = \mathbf{x}_b^{i+1}(n+1) - \mathbf{x}_b^{i+1}(n)$ evolve into analysis differences $\varepsilon_a^{i+1} = \mathbf{x}_a^{i+1}(n+1) - \mathbf{x}_a^{i+1}(n)$, according to the following equation:

$$\varepsilon_a^{i+1} = \varepsilon_b^{i+1} + \mathbf{K}(\varepsilon_o^{i+1} - \mathbf{H}\varepsilon_b^{i+1}) = \mathbf{K}\varepsilon_o^{i+1} + (\mathbf{I} - \mathbf{KH})\varepsilon_b^{i+1} \tag{2.4}$$

where $\varepsilon_o^{i+1}$ are observation differences between two consecutive ensemble members ($\varepsilon_o^{i+1} = \mathbf{y}^{i+1}(n+1) - \mathbf{y}^{i+1}(n)$) that simulate the observation errors $e_o^{i+1}$.

In addition, from one ensemble of analyses one can obtain an ensemble of (perturbed) six-hour forecasts, by integrating the NWP model. The background field can be expressed as

$$\mathbf{x}_b^{i+1}(n) = \mathbf{M}\mathbf{x}_a^i(n) + \vartheta^i(n) \tag{2.5}$$

where $\mathbf{M}$ is the 6h forecast operator and $\vartheta(n)$ are some model perturbations that can be added after the model integration, in order to simulate the model error. These random realizations $\vartheta(n)$ can be built from an estimate of the model error covariance matrix (Mitchell *et al.*, 2002).

It is also possible to obtain an ensemble of background fields by using a different NWP model, $\mathbf{M}_n$, for each ensemble member:

$$\mathbf{x}_b^{i+1}(n) = \mathbf{M}_n\mathbf{x}_a^i(n) \tag{2.6}$$

This approach has been used by Houtekamer *et al.* (1996) in order to simulate the model error.

Finally, the evolution of $\varepsilon_a^i$ into the background differences $\varepsilon_b^{i+1}$ corresponds to the following equation:

$$\varepsilon_b^{i+1} = \mathbf{M}\varepsilon_a^i + \varepsilon_m^{i+1} \tag{2.7}$$

where $\varepsilon_m^{i+1}$ represents the simulated model error, which can be estimated by using different forecast models, as in (2.6), or by adding some model perturbations (according to equation 2.5).

It is important to notice that the basic form of the two equations (2.4) and (2.7), and the involved operators are the same as in the exact error evolution (see equations 2.2 and 2.1). Thus, the ensemble approach has the advantage of simulating two basic and important components of the exact error evolution: the forecast evolution provided by the model and the analysis step

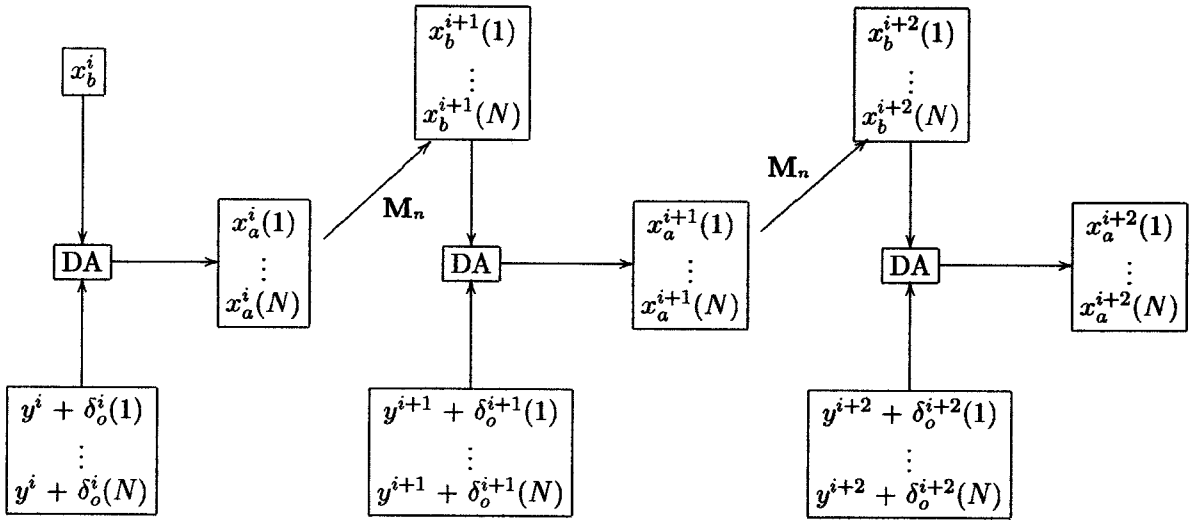


Figure 2.1: Schematic illustration of the Ensemble method, where DA stands for data assimilation. For the first analysis step, the operational (unperturbed) background is used, for all ensemble members.

(represented by $\mathbf{K}$ and $\mathbf{I} - \mathbf{KH}$, which are the respective analysis weights of the observations and of the background). In practice, these two components are simulated by applying a succession of forecasts and analyses to some perturbed states.

In the case of a perfect model framework, it is assumed that the background errors are only due to the time evolution of the analysis errors provided by a perfect model. Thus, $\varepsilon_m$ is simply set to zero in equation (2.7). This means that no model perturbations are added to the 6h-forecasts and that the forecast model used in assimilation experiments is the same for all the ensemble members (i.e., $\mathbf{M}_n = \mathbf{M}$, for $n = 1, ..., N$). This approach is expected to lead, e.g., to an underestimation of the error variances. Nevertheless, it is a simpler approach for a first set of experiments and it also gives the possibility to assess the impact of model error simulations.

## 2.4 The NMC simulation of the error evolution

The NMC method may also be seen as a technique that simulates the evolution of the model state errors. In addition, this approach relies on the assumption that the background error can be estimated by computing the differences between forecasts of different ranges, but valid at the same time. In the present work, differences between 36h and 12h forecasts are used (see figure 2.2). It may be noted that there is a 24h period between the respective starting dates of these two forecasts. During this period, there is a succession of analyses and 6h forecasts. After this 24h period, there is a final forecast step (over a 12 h period).

In order to show the main differences between the NMC approach, the exact and the ensemble error evolutions, it is useful to write the equations for the first and for the final analysis steps in the NMC method (more details are available in Berre *et al.* 2006).

For the first analysis step of the NMC method, the simulated analysis error $\tilde{\epsilon}_a^{i+1}$ is simply equal to the analysis increment:

$$\tilde{\epsilon}_a^{i+1} = \mathbf{K}(e_o^{i+1} - \mathbf{H}e_b^{i+1}) = \mathbf{K}e_o^{i+1} - \mathbf{K}\,\mathbf{H}e_b^{i+1} \tag{2.8}$$

because $\tilde{\epsilon}_a^{i+1} = \mathbf{x}_a^{i+1} - \mathbf{M}\,\mathbf{x}_a^i = \mathbf{x}_a^{i+1} - \mathbf{x}_b^{i+1} = dx^{i+1} = \mathbf{K}(\mathbf{y}^{i+1} - \mathbf{H}\mathbf{x}_b^{i+1})$.

Comparing equations (2.8) and (2.2) it is possible to summarize important features of the NMC method.

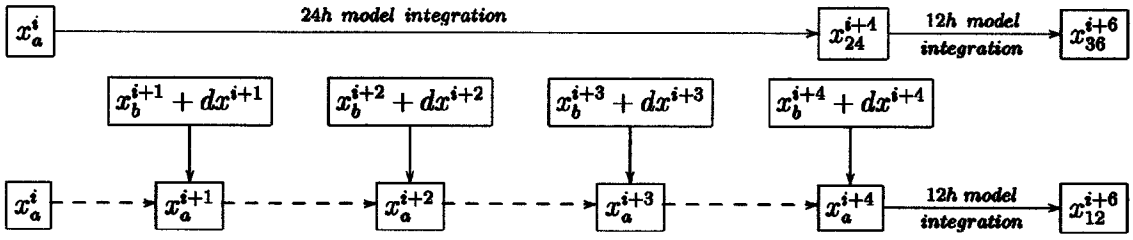


Figure 2.2: Schematic illustration of the NMC method. Here, $\mathbf{x}_a^{i+1}$ and $dx^{i+1}$ are respectively the analysis and the analysis increment at time $t_i$. $\mathbf{x}_{24}^{i+4}$ is the 24h forecast that is valid at time $t_{i+4}$. Similarly, $\mathbf{x}_{36}^{i+6}$ and $\mathbf{x}_{12}^{i+6}$ are respectively the 36h and 12h forecasts that are valid at time $t_{i+6}$.

Regarding the estimated *variances* of the analysis errors, in a similar way as discussed in Bouttier (1994), one may consider the following three limit cases:

- In regions where the data density is rich, and the observation errors have similar variances and similar spatial correlations as the background errors, **KH** and **K** are close to $I/2$. This implies that $e_a^{i+1} \approx \frac{1}{2}(e_o^{i+1} + e_b^{i+1})$, while $\tilde{\epsilon}_a^{i+1} \approx \frac{1}{2}(e_o^{i+1} - e_b^{i+1})$. The consequence is that, in this case, the variances of the analysis increments are similar to the variances of the exact analysis errors. (Note that the variances of $e_o$ and $e_b$ are added, rather than subtracted: $\langle(\tilde{\epsilon}_a^{i+1})^2\rangle \approx \frac{1}{4}(\,\langle(e_o^{i+1})^2\rangle + \langle(e_b^{i+1})^2\rangle\,)$.

- In regions where the data density is rich and the observations have a high quality, **KH** and **K** are close to $I$. This implies that $e_a^{i+1} \approx e_o^{i+1}$, while $\tilde{\epsilon}_a^{i+1} \approx e_o^{i+1} - e_b^{i+1}$. In other words, the variances of the analysis increments are likely to be an overestimation of the variances of the exact analysis errors.

- In regions where the data density is poor and the (few) observations have a low quality, **KH** and **K** are close to 0. This implies that $e_a^{i+1} \approx e_b^{i+1}$, while $\tilde{\epsilon}_a^{i+1} \approx 0$. In other words, in data sparse areas, the simulated analysis error variances are likely to be underestimated by the NMC approach.

Regarding the estimated *correlations* of the analysis errors, one may note that the exact operator $(\mathbf{I} - \mathbf{KH})$ is replaced by **KH** in equation (2.8). If the observation error covariances are similar to the background error covariances (as evoked in the first limit case above), these two operators are similar, and the analysis error correlations can be expected to be well approximated. However, the observation errors are usually less spatially correlated than the background errors. As discussed in section 1.3.4, following Daley (1991, section 4.5), this implies that $(\mathbf{I} - \mathbf{KH})$ acts as a high-pass filter, while **KH** is rather a low-pass filter. The consequence is that in this case the analysis error correlations are likely to be overestimated by the NMC method. In other words, the NMC method tends to overestimate the correlation length scale of the analysis errors.

For the "final" analysis step (which corresponds to the 4th analysis step in the 24 h period) in the NMC method, the simulated analysis error $\tilde{\epsilon}_a^{i+4} = \mathbf{x}_a^{i+4} - \mathbf{x}_{24}^{i+4}$ is:

$$\tilde{\epsilon}_a^{i+4} = \tilde{\epsilon}_b^{i+4} + dx^{i+4} = (\sum_{k=1}^{3} \mathbf{M}^{4-k}\, dx^{i+k}\,) + dx^{i+4} \tag{2.9}$$

where $\tilde{\epsilon}_b^{i+4}$ can be seen as the NMC background perturbation, which contains the contributions of the earlier analysis increments, that were evolved in time and accumulated.

Thus, in the NMC method, the analysis step consists in adding the (unperturbed) analysis increment to a background perturbation (which corresponds itself to some earlier evolved increments). In contrast, in the exact and ensemble evolutions, the analysis step consists in applying the analysis equation to a vector of background differences and to a vector of observation differences (the differences are the "true errors" in the exact evolution).

The last equation can be re-written as:

$$\tilde{\epsilon}_a^{i+4} = \sum_{k=1}^{4} \mathbf{M}^{4-k}\, dx^{i+k}$$

In addition, in the NMC method, this final analysis perturbation is evolved during 12 additional hours, which provides the final 36h-12h forecast differences, $\tilde{\epsilon}_b^{i+6} = \mathbf{x}_{12}^{i+6} - \mathbf{x}_{36}^{i+6}$ (that intend to simulate $e_b^{i+6}$):

$$\tilde{\epsilon}_b^{i+6} = \mathbf{M}^2\, \tilde{\epsilon}_a^{i+4} = \mathbf{M}^2\, (\sum_{k=1}^{4} \mathbf{M}^{4-k}\, dx^{i+k})$$

which can be compared with the corresponding equation for the background differences in the ensemble method:

$$\varepsilon_b^{i+6} = \mathbf{M}\, \varepsilon_a^{i+5}$$

The comparison between the last two equations allows to summarize the main three characteristics of the NMC method, compared with the ensemble method:

- the involvement of longer forecast ranges (see the occurrence of the matrices $\mathbf{M}^2, \mathbf{M}^{4-i}$ instead of the 6h matrix $\mathbf{M}$) ;
- the accumulation of several increments (see the occurrence of the operator $\sum$) ;
- the involvement of analysis increments $dx$, instead of analysis differences $\varepsilon_a$.

The involvement of longer forecast ranges has been evoked by Fisher (2003), to explain the differences between the NMC and ensemble methods. It will be shown in section 3.3.3 that the differences in the representation of the analysis step (i.e. the third item in the list above) also play an important role.

# Chapter 3

# The background error covariances estimated by an ensemble approach in a perfect-model framework

*This chapter corresponds essentially to the experimental part of Belo Pereira and Berre (2006). It has been extended by including some additional experimental results.*

## 3.1 Introduction

The diagnosis of the background error covariances of the Arpège global NWP model, estimated by the ensemble approach using a perfect-model assumption, is presented in this chapter. Formally, this assumption corresponds to setting $M_n = M$ (or equivalently $\vartheta^i(n) = 0$), for any ensemble member $n$ (see section 2.3). The resulting covariance estimates have been compared with those that were previously operational, by performing both diagnosis and impact studies: the global (i.e. spatially averaged) ensemble covariances have been compared with those of the NMC method, and the local (i.e. spatially varying) ensemble variances have also been compared with the local operational variances (the latter correspond to a static 3D map of variances, that reflects some typical data density contrasts).

Some emphasis has been given to study the role of the analysis step in the simulation of the background errors. Moreover, the background and the analysis errors estimated by the ensemble approach have been compared.

Significant effort has been devoted to the diagnosis and examination of the local ensemble correlations. This has been done by developing an economical method to estimate the local correlation length scale. The study of these local correlations gives interesting information about the existing heterogeneities and anisotropies. This can also be useful for the future evaluation of heterogeneous covariance formulations, such as those based on wavelets (Fisher (2003), Deckmyn and Berre (2005)) or recursive filters (Wu *et al.* 2002).

The experimental framework is described in section 3.2. Section 3.3 concerns the comparison between the global covariance estimates of the ensemble and NMC methods. The local covariance estimates are diagnosed in section 3.4. The impact studies of the ensemble global covariances and of the ensemble local variances are then described in section 3.5.

## 3.2 The experimental framework

The results shown in the present chapter were produced from one ensemble with five 4D-Var assimilation cycles, that were performed from the 1st of February to the 24th of March 2002. The first analysis ensemble is created by randomly perturbing the observations, and by using the (unperturbed) operational background field. The model integration will then provide an ensemble of backgrounds (see figure 2.1). Thus, after the first analysis step, the perturbed analyses result from both observation and background perturbations. The amplitude of the background differences grows thus from zero to stable values during the first three days. This preliminary period is therefore not included in the statistics, which are rather calculated from 04/02 to 24/03 (which is a 49 day period).

The members were arbitrarily numbered from 1 to 5 $(x_b(1), x_b(2), x_b(3), x_b(4), x_b(5))$, from which four sets of member differences were calculated $(x_b(1) - x_b(2), x_b(2) - x_b(3), x_b(3) - x_b(4), x_b(4) - x_b(5))$. So, the correlations and standard deviations of the background error are computed from $4 \times 49 = 196$ differences between 6h forecasts, for each of the four daily analysis times (00, 06, 12 and 18 UTC). Mostly, the results from fields valid at 18UTC will be shown.

It may be mentioned that using differences between background states from separate perturbed experiments is equivalent to the use of differences between a perturbed background and the background from an unperturbed control experiment, except for a factor 2 (see e.g. Berre *et al.* (2006): it can be shown that the covariance of the difference between two perturbed backgrounds is equal to twice the covariance of the single background perturbations).

The experiments have been performed with the Arpège global model (in its non-stretched version, i.e., with a uniform resolution), and its 4D-Var scheme (Rabier *et al.* (2000); Veersé and Thépaut (1998)). The model has been integrated at T299 triangular truncation with 41 levels. The formulation of the **B** matrix that is used in the Arpège 4D-Var is described in appendix B, following the description given by Derber and Bouttier (1999).

## 3.3 Global background error covariances

This section presents the comparison between the ensemble and NMC methods, relatively to the estimation of the global background error covariances. The covariances were computed in spectral space at truncation T179 and were examined for the following variables: vorticity ($\zeta$), divergence ($\eta$), temperature ($T$), logarithm of surface pressure ($P_s$) (which will be referred to as surface pressure in the remainder of the text), and specific humidity ($q$).

The results are focused on the vertical profiles of standard deviation, on the horizontal correlation spectra, and on the horizontal correlation length scales. The influence of the analysis step on these differences will be studied in the final part of this section.

### 3.3.1 Standard deviations

The vertical profiles of the total standard deviations are used to determine how the (horizontally averaged) amplitudes of the background errors vary as a function of height.

Figure 3.1 presents the standard deviations of the temperature and vorticity background errors, estimated by the NMC and ensemble methods. One can notice that there are some similarities between the two methods. For instance, both methods indicate that the background errors of temperature show small variations in vertical, in particular in troposphere. Another common feature is that the background errors of vorticity increase significantly with height, reaching their maximum near the tropopause. This maximum is related to the jet streams in

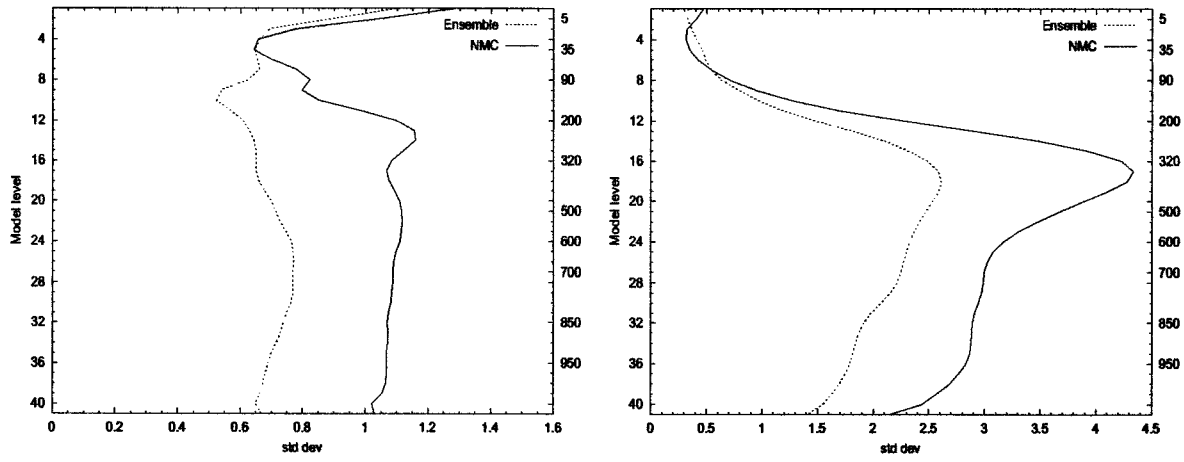


Figure 3.1: Vertical profile of standard deviation of the temperature (K) (left side) and vorticity ($10^{-5}s^{-1}$) (right side) background errors. In the right axes is plotted the atmospheric pressure (in hPa) corresponding to a certain model level for a standard atmosphere.

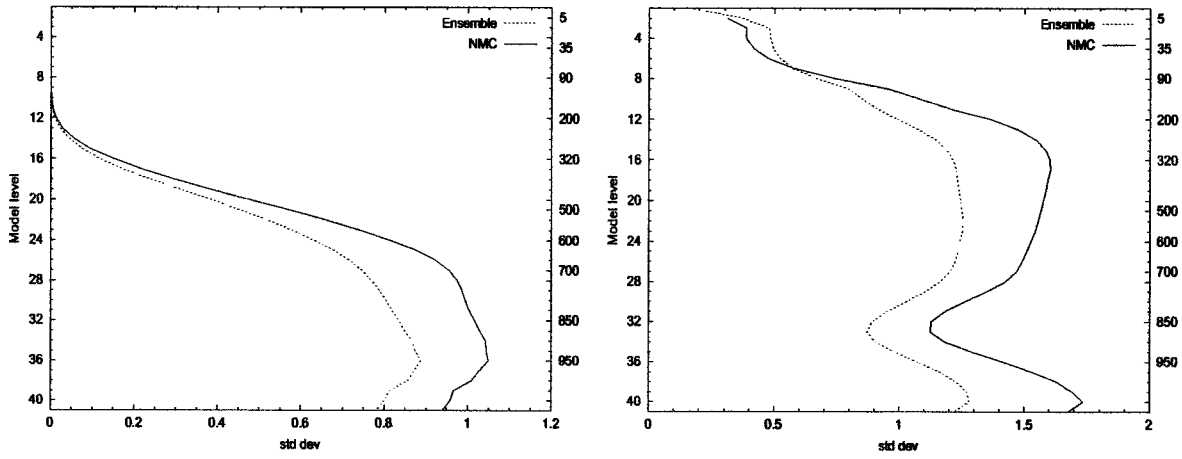


Figure 3.2: Vertical profile of standard deviation of the specific humidity ($g\ kg^{-1}$) (left side) and divergence ($10^{-5}s^{-1}$) (right side) and background errors.

the middle latitudes. Nevertheless, the contrast between the jet level and the low levels is weaker in the ensemble method than in the NMC method.

The standard deviations are obviously larger in the NMC method than in the ensemble method. The absence of model error simulation in the current version of the ensemble experiments is likely to contribute to this difference. Moreover, the accumulation of four analysis increments and the involvement of long forecast ranges in the NMC method contribute probably also to this result.

According to both methods, the largest background errors of $q$ occur in the lower troposphere, below 700hPa, where more specific humidity is available. Above this level, the background errors of $q$ decrease drastically with height. For divergence, the largest background errors are found near the surface and in the upper and middle troposphere, for both methods (figure 3.2). Furthermore, for these variables, the standard deviations estimated by the NMC method are also larger than those from the ensemble method.

### 3.3.2 Horizontal correlations

The correlation spectra can be deduced from covariance spectra by normalizing it by the global variance. Therefore, the correlation spectra will be shown instead of the covariance spectra, since this allows to compare more directly the relative contributions of the different horizontal scales, in the two methods (see section 1.3.3).

According to the meteorological nomenclature, the wave number, $k$, indicates the number of waves along the latitude circle, i.e., $k = 2R\pi cos\phi/L$, where R is the earth radius, $L$ the wavelength and $\phi$ is the latitude. Thus, the wave number 1 corresponds to a wavelength equal to the earth's circumference, at the equator. Following this nomenclature, the scales resolved the Global model ARPEGE may be grouped into three categories based on the wave number $k$; large or planetary scales associated to long waves for $k$ between 1 and 5 (ultralong if $k < 3$); synoptic scales for $k$ between 6 and 20; and mesoscales for $k \geq 21$.

The figure 3.3 presents the variation with height of the autocorrelation spectra of the temperature and vorticity background error estimated by the ensemble method. From this figure it is clear that the maximum variance[1] of the background error is shifted towards the larger scales with height. In other words, the contributions from the mesoscale phenomena to the background errors are emphasized in the lower troposphere and the contributions from large scales are mostly important in the upper troposphere and in stratosphere.

Moreover, the results show that the largest contributions to the vorticity and divergence background errors are due to the mesoscale phenomena, while for temperature and specific humidity the largest contributions are associated to the synoptic scales. This is illustrated for vorticity and temperature in figure 3.3.

The figure 3.4 presents the correlation spectra for temperature and vorticity background errors (at model level 21, near 500hPa). One can see that compared with the NMC method, the ensemble approach emphasizes the relative contributions of the small scales (i.e. compared with the contributions of the large scales). This result is also found for the other variables. As illustrated in figure 3.5, this implies that the corresponding correlation length scales (computed from equation 1.48, section 1.3.3, and defined by equation 1.42, section 1.3.2) are smaller in the ensemble than in the NMC method.

The increase of length scale with height, found in other studies (e.g. Rabier *et al.*, 1998), can also be identified in figure 3.5, in accordance with the increase with height of the contributions of large scale phenomena. Moreover, according to both methods, the length scale of vorticity is smaller than the one of temperature. This is consistent with the larger importance of mesoscale phenomena for the vorticity background error than for the temperature one, as showed in figure 3.3.

The enhancement of the small scale horizontal contributions implies also that the vertical correlations are sharper in the ensemble method (see section 3.4.4). Another consequence is that the contributions of the unbalanced components are emphasized in the multivariate formulation. In other words, the amount of explained variance (equation B.12, appendix B) by the balance relationships is smaller in the ensemble method than in the NMC method. This is illustrated in figure 3.6 for temperature. All these results are consistent with those mentioned by Fisher (2003).

### 3.3.3 The influence of the analysis step

In sections 2.3 and 2.4, it has been demonstrated that the ensemble and NMC methods differ with respect to the representation of the analysis step. The ensemble method simulates the reduction of the model state errors when combining the (uncertain) background with the (uncertain) observations. To contrast with this, the NMC method is rather relying on the accumulation and time evolution of four (successive) analysis increments.

[1] normalized variance

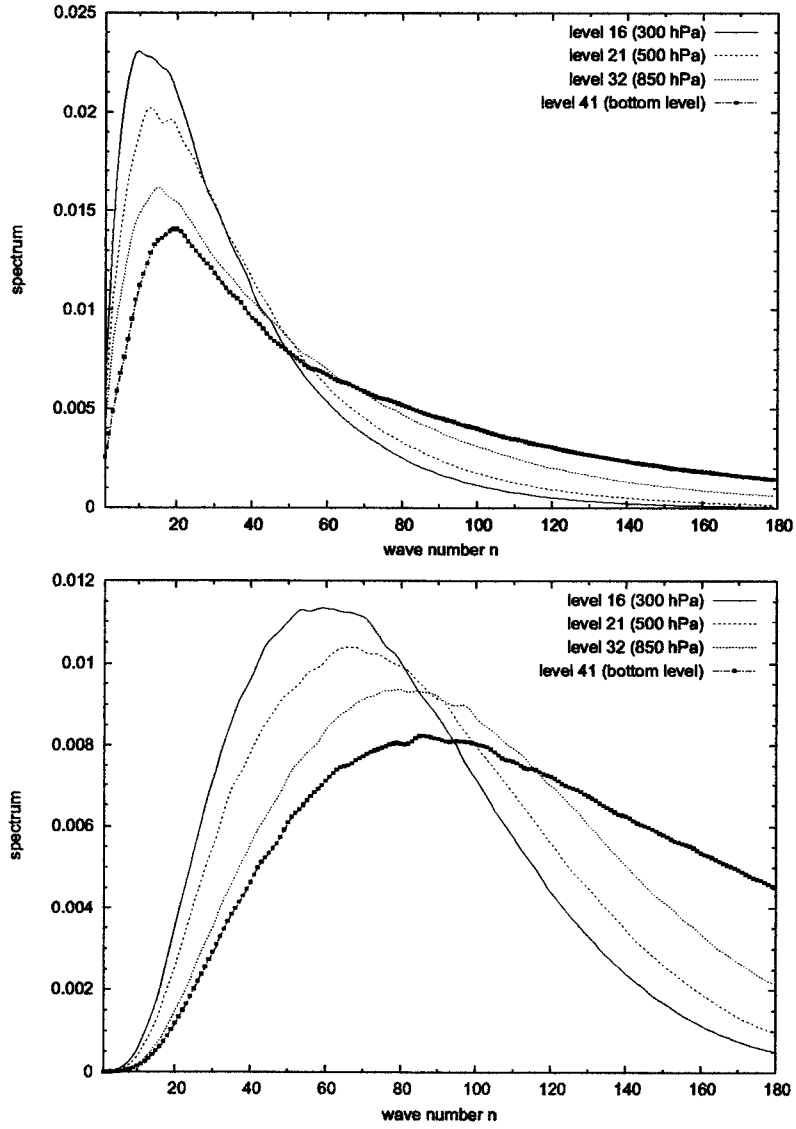


Figure 3.3: Autocorrelation spectra of the background error of temperature and vorticity for different levels, estimated by the ensemble method.

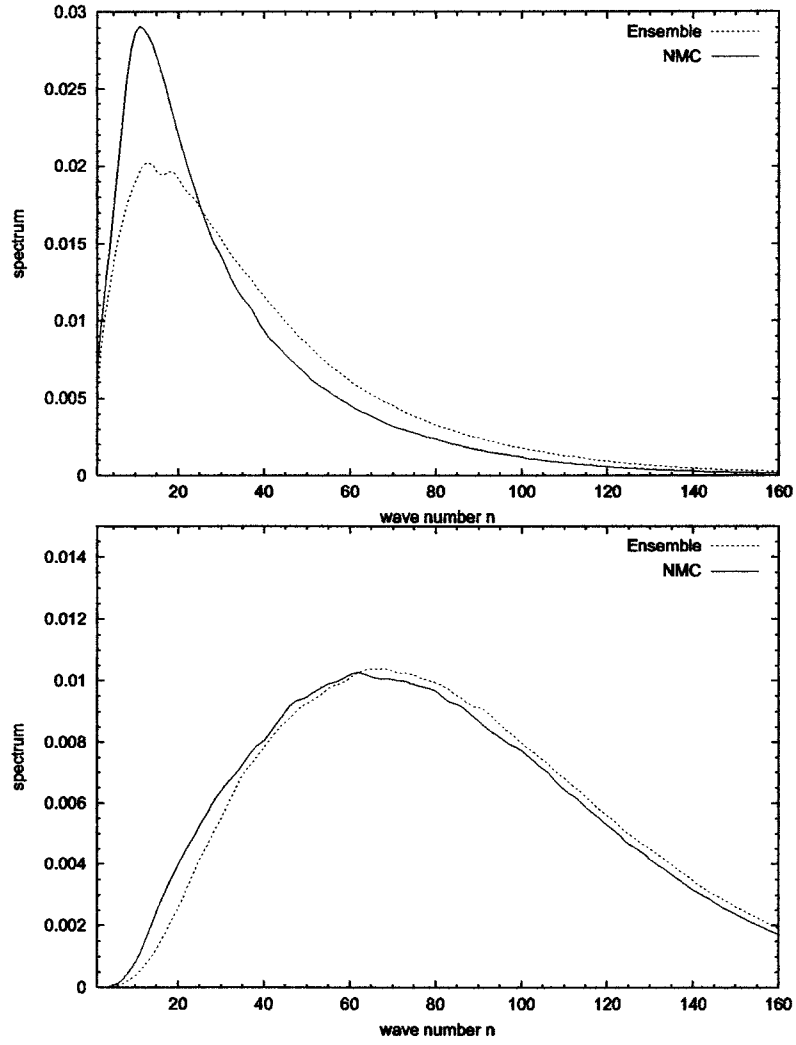


Figure 3.4: Autocorrelation spectra of the background error of temperature (top panel) and vorticity (bottom panel) at level 21 (near 500hPa) for NMC (full line) and ensemble (dotted line) methods.

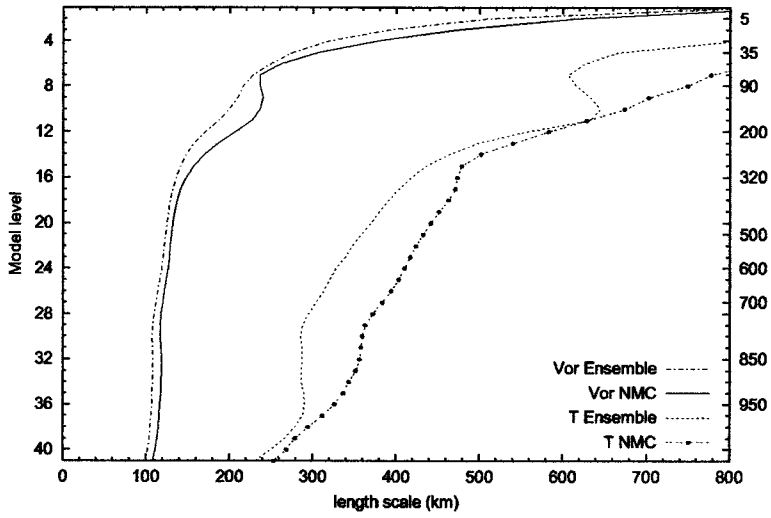


Figure 3.5: Horizontal length scale of the autocorrelation function of the vorticity and temperature background error, estimated by the NMC and ensemble method.

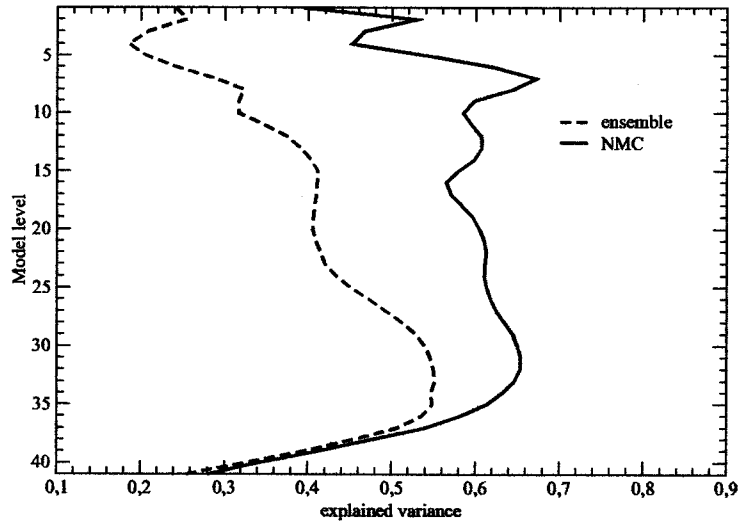


Figure 3.6: Explained variance for the temperature background error, estimated by the NMC and ensemble methods. This is a measure of the strength of the mass/wind balance (couplings between temperature, vorticity and divergence) in the background errors (see appendix B, equation B.12).

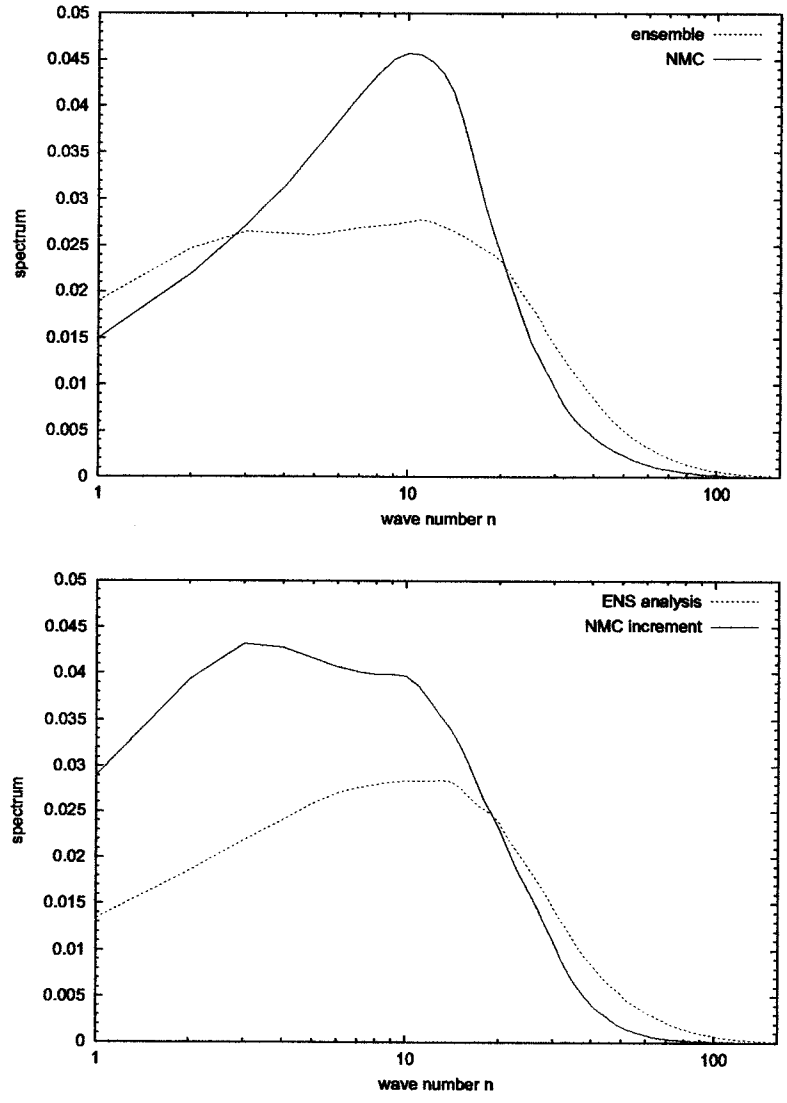


Figure 3.7: Autocorrelation spectra for surface pressure. Top panel: Background error spectra estimated by the ensemble (dashed line) and NMC (solid line) methods. Bottom panel: spectra of the analysis dispersion in the ensemble method (dashed line) and of the analysis increment in the NMC method (solid line).

An indication of the influence of these analysis representation differences is shown in figure 3.7. The top panel represents the background error correlation spectra of $P_s$, estimated by the ensemble and NMC methods. The contributions of the small scales and of the planetary scales are emphasized in the ensemble method.

The bottom panel corresponds to the comparison between the correlation spectra of the analysis increment (in the NMC method) and of the analysis dispersion (in the ensemble method). It appears that the analysis increment spectrum is much larger scale than the analysis dispersion spectrum. Such a difference is consistent with the expectation that the correlation length scales of the analysis increments are larger than the correlation length scales of the analysis errors (see section 2.4). As the analysis increment is one fundamental ingredient of the NMC method, this contrast explains the relative larger contributions of the small scales to the background errors in the ensemble method. This indicates that the analysis representation differences play an important role in the scale differences between the two methods.

## 3.4 Local background error covariances

In this section, the geographical variations of the ensemble covariances are studied. Firstly, the ensemble local standard deviations will be compared with the operational local standard deviations, and also with the NMC local standard deviations.

Secondly, the geographical variations of the ensemble correlations will be examined. This will be done in particular by introducing an economical estimation of the local correlation length scales. Concerning this aspect, a comparison will be done with the NMC method.

Finally, the corresponding differences in the vertical correlations and in the analysis/background errors will be also illustrated.

### 3.4.1 Standard deviations

All the statistics presented in this subsection (as in the whole section) were computed on a $1.5° \times 1.5°$ latitude-longitude grid (which is comparable to a truncation T130 approximately). However, the standard deviations maps presented here have been truncated at T21 for vorticity and at T79 for temperature. These truncations have been chosen in order to filter the sampling noise, while retaining the relevant features.

The figure 3.8 shows the maps of normalized[2] standard deviations of the vorticity background error at 500hPa, estimated by the ensemble method and used in the operational experiments. The normalized standard deviations correspond to some horizontal modulations, that are applied to the global covariances (see section 3.3.1).

Some common features between the three estimates can be identified regarding the latitudinal variations: the standard deviations are smaller in the tropics than in the middle latitudes. However, these contrasts are strongest in NMC method and smallest in the operational map. The smallest values of standard deviations over the tropics in the NMC method (when compared with the other estimates) are consistent with the fact that the NMC approach tends to underestimate the background errors in data poor areas, for the reasons explained in section 2.4.

Concerning the influence of data density, there are important differences between the NMC method and the other two estimates. From both the ensemble and the operational maps, it is visible that the standard deviation has relatively small values over data rich regions such as Europe and the United States of America (USA), particularly in the operational estimate.

On the other hand, in the estimate from the NMC method, local maxima are visible in data poor areas, such as over Northern Atlantic and North Pacific, as over data rich areas like North America. In addition, the standard deviation has relatively large values over Europe.

[2]the normalization consists simply in a division by the horizontal average of the standard deviation

The differences found between the NMC method and the other approaches are probably related to the tendency of the NMC method to overestimate the background errors in data-rich areas (see section 2.4).

The results show also that the local maxima over the middle and high latitude oceans (such as over the Northern Pacific, the Northern Atlantic, and the circumpolar ocean of the Southern Hemisphere) are strongest in the NMC method and weakest in the operational estimate.

Finally, the ensemble method provides also some local maxima near the North Pole and over the Himalayas, which are not present in the other two maps. These large values in North Pole are consistent with the lack of snow analysis in the ARPEGE model.

The geographical variations of the ensemble standard deviations for temperature background error have been also compared with those provided by the NMC method. The estimates at model levels 37 (near 970hPa) and 27 (near 700hPa) will be used to illustrate some typical differences (figures 3.9 and 3.10, respectively).

Near 970hPa, some similar latitudinal variations are found according to the two methods. For instance, the mid-latitude values are larger than in the tropics, and some large values around the North Pole are visible in the two estimates.

One of the most striking differences is the occurrence of some strong maxima in the central regions of North America, according to the NMC method. This is consistent with the results found for vorticity. However, when the ensemble method is used, the background errors have a local maximum in Canada, near the Hudson bay, but they have relatively small values in USA. A local minimum over Europe is also more visible in the ensemble method than in the NMC method.

Over e.g. Europe and North America, the differences between the two methods can be partly explained by the tendency of the NMC method to overestimate the background errors in data rich areas, as explained in section 2.4. On the other hand, it may be that the perfect model assumption leads to an underestimation of the background errors in some regions by the ensemble method. For instance, one could consider that the temperature errors in the Boundary Layer in mountains areas (e.g. Rocky mountains and Alps) might be underestimated by the ensemble method.

The figure 3.9 also indicates that, according to the ensemble method, large values of background errors occur in Equatorial Africa. This is consistent with the large atmospheric variability observed in this area, associated with the Inter-Tropical Convergence Zone (ITCZ). Such a feature is not well captured by the NMC method. More generally, in the tropics, the magnitude of the background error is larger when estimated by the ensemble method than by the NMC method. These differences may be explained by the fact that the NMC method tends to underestimate the background errors in data sparse areas, as mentioned before.

Near 700hPa, when the background errors are estimated by the ensemble method, the contrast between data rich and data sparse regions is very clear (figure 3.10). At this level, as at other levels, the NMC method is not able to represent such a feature.

Moreover, in the Southern Hemisphere, the largest background errors of the ensemble method are located in the Tropical Eastern Pacific, and in the Tropical Atlantic. In addition, near 850hPa (not shown), according to the ensemble method, very large errors are located in the North Pacific in adjacent areas to the North American west coast. Relatively large values are also found near the West coast of Iberian Peninsula and near the NW and SW coast of Africa. These results are related to the presence of marine stratocumulus, that have a quasi-permanent character in these regions. This may reflect that some large uncertainties exist in these regions: these uncertainties may be related to some strong sensitivity of the shallow convection and radiation parametrizations, with respect to the initial conditions of the NWP model. The NMC method also reflects such uncertainties, but less clearly than the ensemble method. This may be due to an underestimation of analysis errors (by the NMC method) over these data-poor oceanic regions.

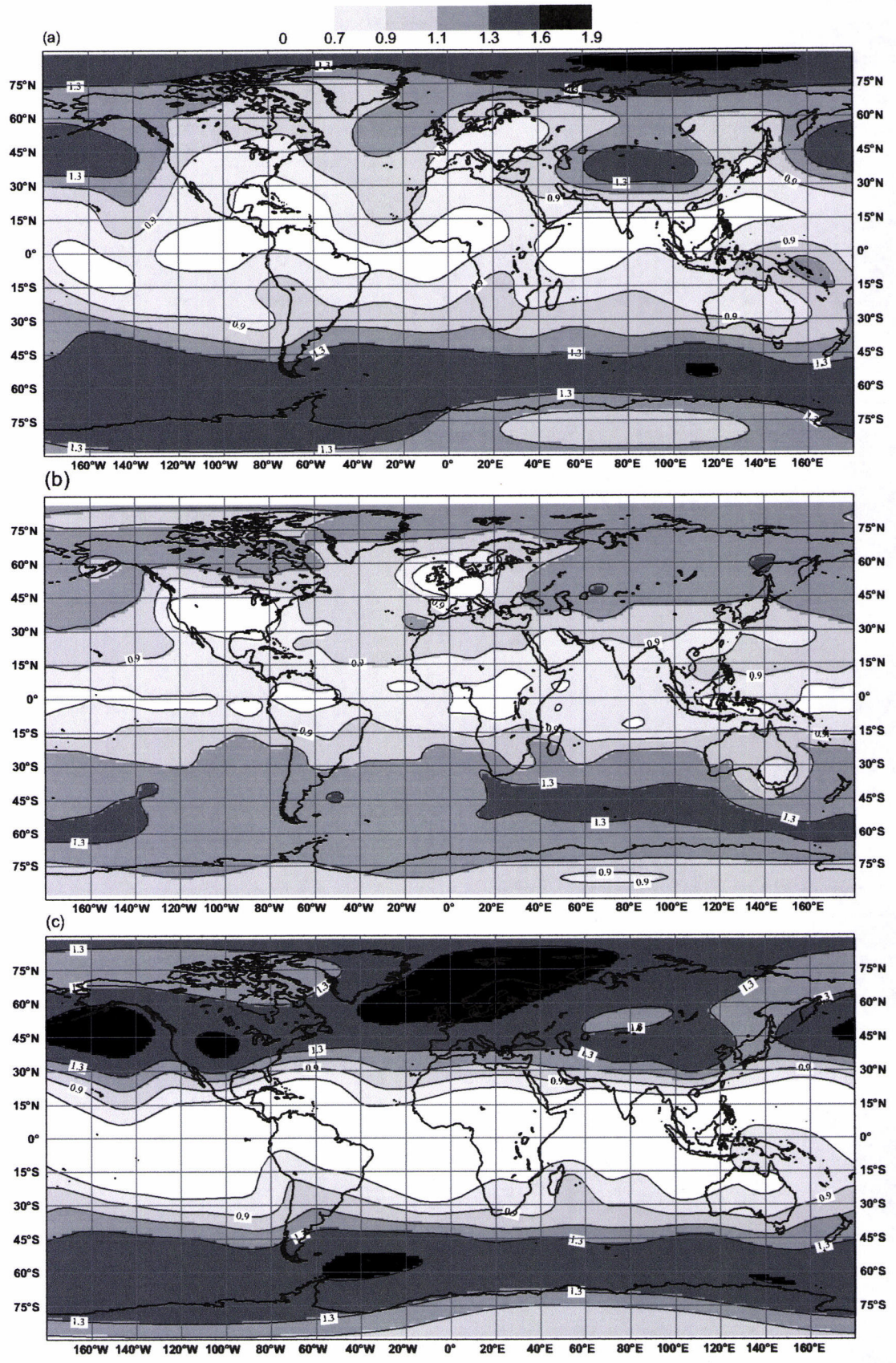


Figure 3.8: Normalized standard deviations of the vorticity background error (around 500 hPa), (a) estimated by the ensemble method, (b) used in the operational experiments and (c) estimated by the NMC method. (The map (b) has truncation T42, as used in the former operational 4D-Var).

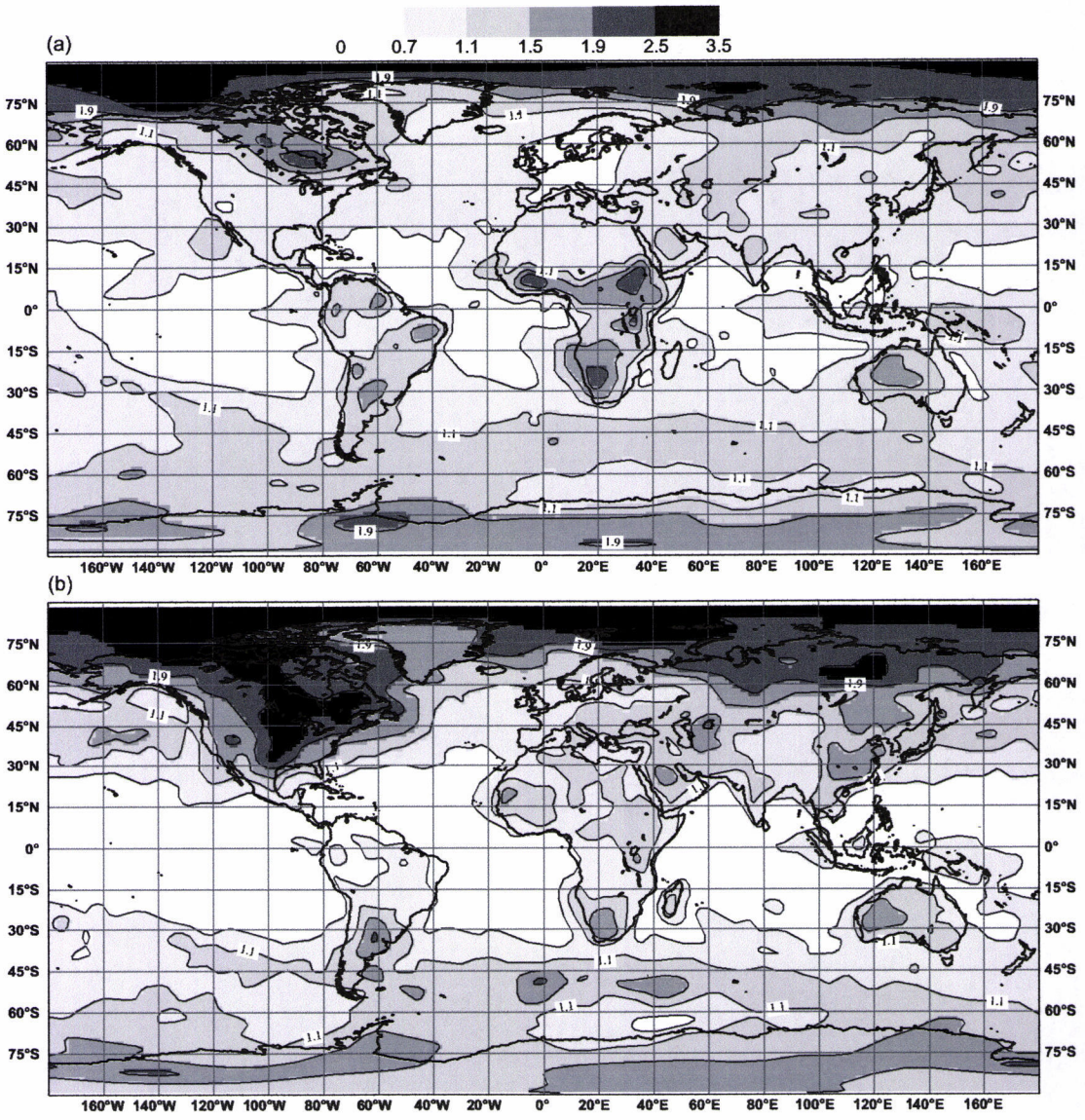


Figure 3.9: Normalized standard deviations of the temperature background error at level 37 (i.e. around 970 hPa), using the ensemble method (top) and the NMC method (bottom).

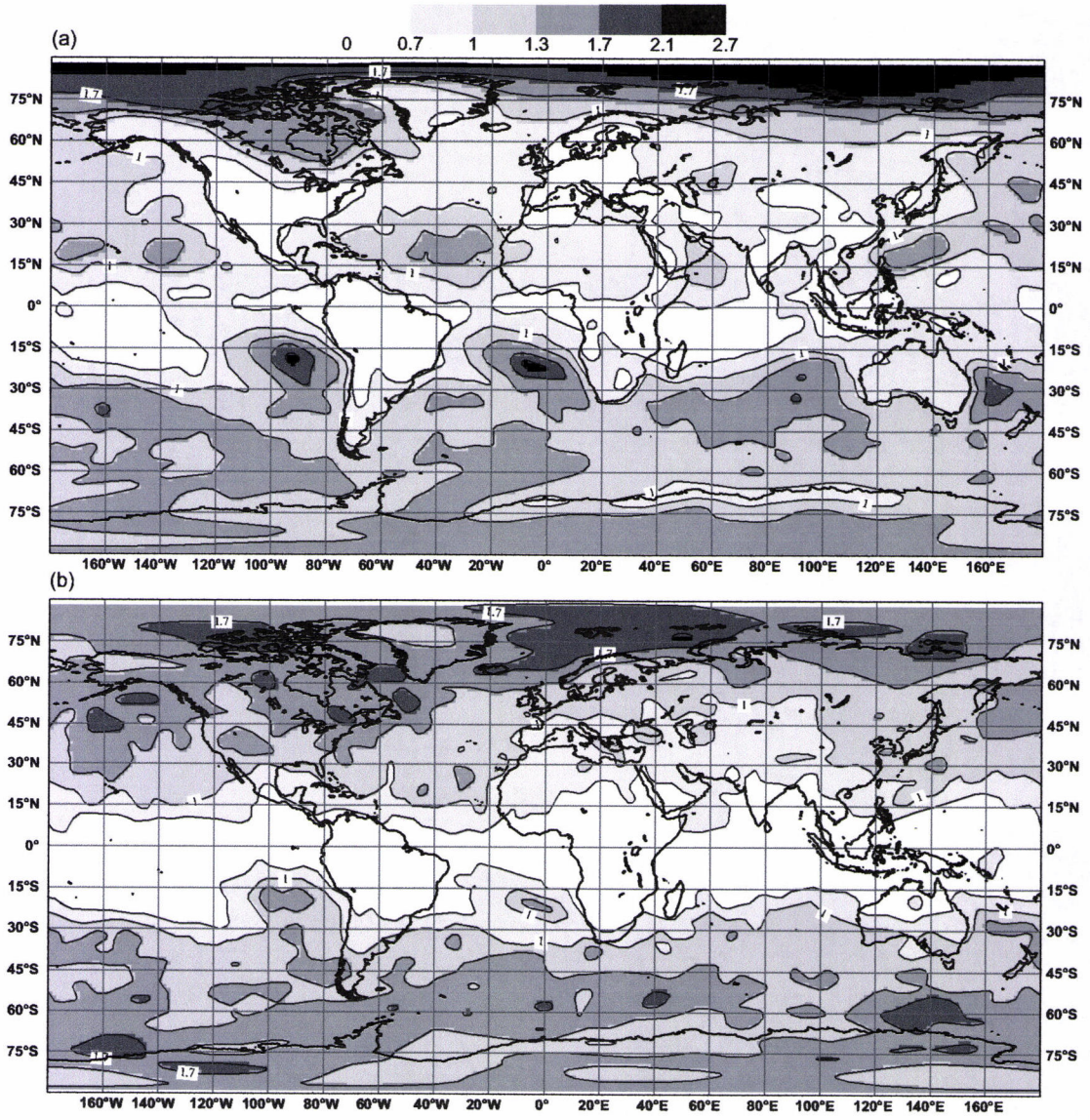


Figure 3.10: Normalized standard deviations of the temperature background error at level 27 (i.e. around 700 hPa), using the ensemble method (top) and the NMC method (bottom).

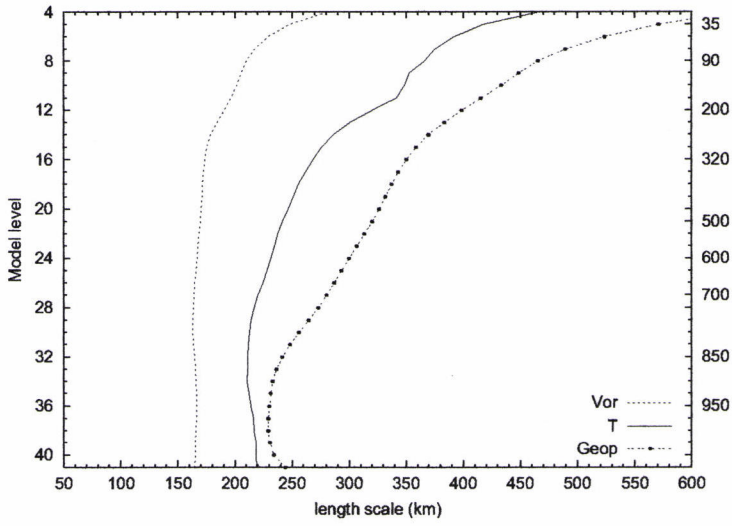


Figure 3.11: Length scale $L$ of background error estimated by the ensemble method in gridpoint space.

### 3.4.2 Horizontal length scales

Diagnosing the full geographical variations of the correlations is a significant challenge: the total number of correlation values to be calculated is the square of the number of gridpoints. This is much larger than what is required for e.g. the calculation and the representation of the geographical maps of standard-deviation. In this context, it is meaningful to introduce an economical estimation of the local correlation length scale. This has been achieved by using the background error variance of each variable and of its derivative. For any error $e$, it is possible to define a zonal length scale ($L_x^e$) and a meridional length scale ($L_y^e$) as (the derivation of these expressions is presented in appendix C):

$$L_x^e = \sqrt{\frac{\sigma^2(e)}{\sigma^2(\frac{\partial e}{\partial x}) - \left(\frac{\partial \sigma(e)}{\partial x}\right)^2}}, \qquad L_y^e = \sqrt{\frac{\sigma^2(e)}{\sigma^2(\frac{\partial e}{\partial y}) - \left(\frac{\partial \sigma(e)}{\partial y}\right)^2}} \tag{3.1}$$

where $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ are respectively the derivatives in the $x$ and $y$ directions, $\sigma^2(e)$ and $\sigma^2(\frac{\partial e}{\partial x})$ are respectively the variance of the error $e$ and of $\frac{\partial e}{\partial x}$. The 2D correlation length scale is given by $L^e = \sqrt{({L_x^e}^2 + {L_y^e}^2)/2}$.

This method of computing $L$ is economical: for a domain with $N$ gridpoints, it is necessary to calculate only $3N$ variances, instead of computing $N^2$ correlations or covariances. Moreover, this approach provides a description of the geographical variations of the correlation functions.

In order to validate this estimate of the length scales, the global length scales have been compared with the ones computed in spectral space. The length scales computed in gridpoint space are shown in figure 3.11 for vorticity, temperature and geopotential. Comparing this figure with figure 3.5 it can be seen that the length scales estimated in gridpoint space are very similar, but larger than those estimated in spectral space, mainly for vorticity, which is consistent with the lower resolution used in gridpoint space than in spectral space. The results also show that $L$ is larger for geopotential than for the other variables. Furthermore, the correlation length scale of geopotential background error increases significantly with height. This is consistent with the fact that the atmospheric waves which predominate in the stratosphere and in the upper troposphere are longer than those in the lower troposphere.

The length scale estimates that are provided by the ensemble and NMC methods will be noted $L^e_{ENS}$ and $L^e_{NMC}$ respectively. The latitudinal variations of $L^e_{NMC}$ and $L^e_{ENS}$ were examined, for

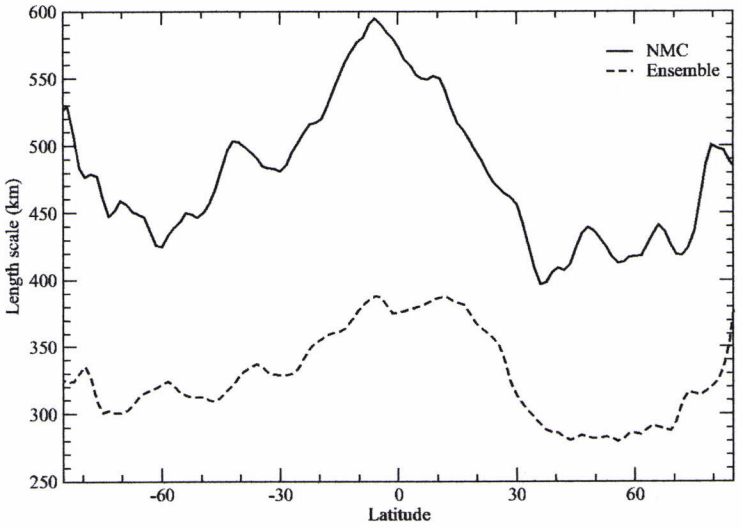


Figure 3.12: North-south variation of the length scale of surface pressure, estimated by the ensemble (dashed line) and NMC (full line) methods.

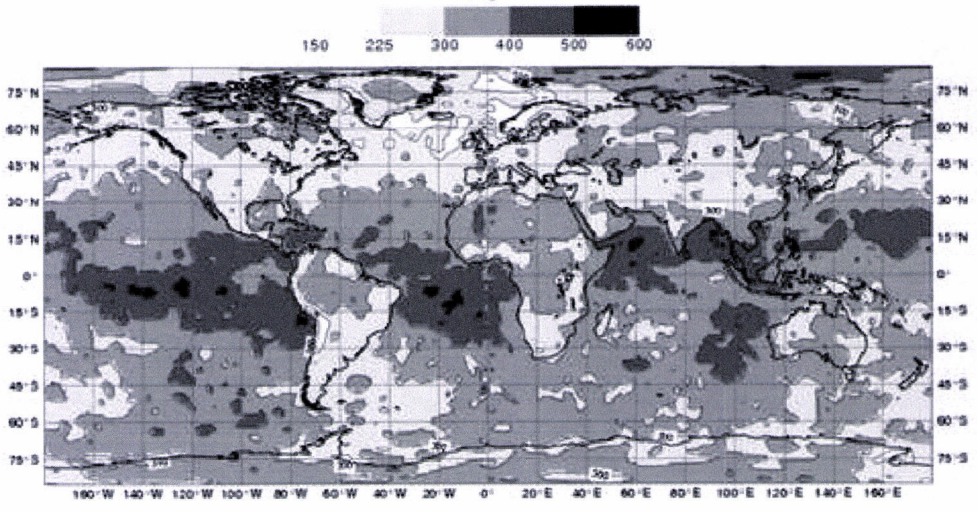


Figure 3.13: Geographical variation of the length scale of surface pressure, estimated by the ensemble method.

Table 3.1: Regions used to study the heterogeneities of the background error statistics and to compute the forecast scores (as it will be shown in section 3.5).

| Acronym | Region | latitudes | longitudes |
|---|---|---|---|
| NORD20 | Northern Hemisphere | 20°N - 90°N | |
| EUROPE | Europe (over land) | 30°N - 70°N | 10°W - 35°E |
| EURATL | Europe and East region of North Atlantic | 30°N - 70°N | 10°W - 35°E |
| NORAMER | North America (over land) | 25°N - 60°N | 145°W - 50°W |
| AMNORD | North America and surrounding ocean areas | 25°N - 60°N | 145°W - 50°W |
| ASIE | Asie | 25°N - 65°N | 60°E - 145°E |
| TROPIQ | tropics | 20°S - 20°N | |
| SUD20 | Southern Hemisphere | 20°S - 90°S | |
| AUS/NZ | Australia/New Zealand and surrounding ocean areas | 10°S - 55°S | 90°E - 160°E |

different meteorological variables. According to both methods, the correlation length scale is larger in the tropics than in the middle and high latitudes, for all the variables. This is illustrated for surface pressure in figures 3.12 (for both methods) and 3.13 (for ensemble method). Following the physical explanation from Lindzen and Fox-Rabinovitz (1989), Ingleby (2001) suggested that the increase of the horizontal correlation length scale, when approaching the equator, reflects the latitudinal dependence of the Rossby radius of deformation. From figure 3.13 it is also obvious that length scales are larger over sea than over land.

From figure 3.12, it is also clear that $L^e_{ENS}$ is much shorter than $L^e_{NMC}$. Moreover, in the middle latitudes, the correlation length scale of $P_s$ is smaller in the Northern Hemisphere than in the Southern Hemisphere. This is partially related to the larger presence of land surfaces in the Northern Hemisphere.

The length scales for different regions of the globe were also examined. The corresponding vertical profiles of length scale are shown for temperature in figure 3.14. These regions are defined in table 3.1.

According to the ensemble method, the length scale is largest in the tropics and smallest in the data rich regions (for instance, in EUROPE and NORAMER areas). This result is in agreement with other studies concerning the effects of latitude on the length scales (e.g. Lindzen and Fox-Rabinovitz (1989), Ingleby (2001)) and the effects of data density (Bouttier, 1994).

In contrast, between 950hPa and 500hPa, the largest $L^e_{NMC}$ values are found in the NORAMER region, rather than in the TROPIQ area. This is mostly related to the largest differences between $L^e_{NMC}$ and $L^e_{ENS}$ over NORAMER. This indicates that the background correlation overestimation is more pronounced over data rich areas. Moreover, contrarily to the ensemble method, it was found that in the middle troposphere, the $L^e_{NMC}$ value for temperature is larger over land than over sea (not shown).

The figure 3.15 shows the comparison between $L^e_{ENS}$ for the analysis error and length scales of $\tilde{\epsilon}_a^{i+4}$, of $\tilde{\epsilon}_b^{i+6}$ (estimated by the NMC method) and of the analysis increments ($dx$) for temperature over different regions. The most striking feature in this figure is the large difference between the correlation length scales of the analysis error estimated by the ensemble method and those from the analysis increments, particularly in data sparse regions, such as the SUD20 and the TROPIQ domains.

This result is consistent with results presented in section 3.3.3 and with the theoretical considerations claiming that the correlation length scales of the analysis increments are larger

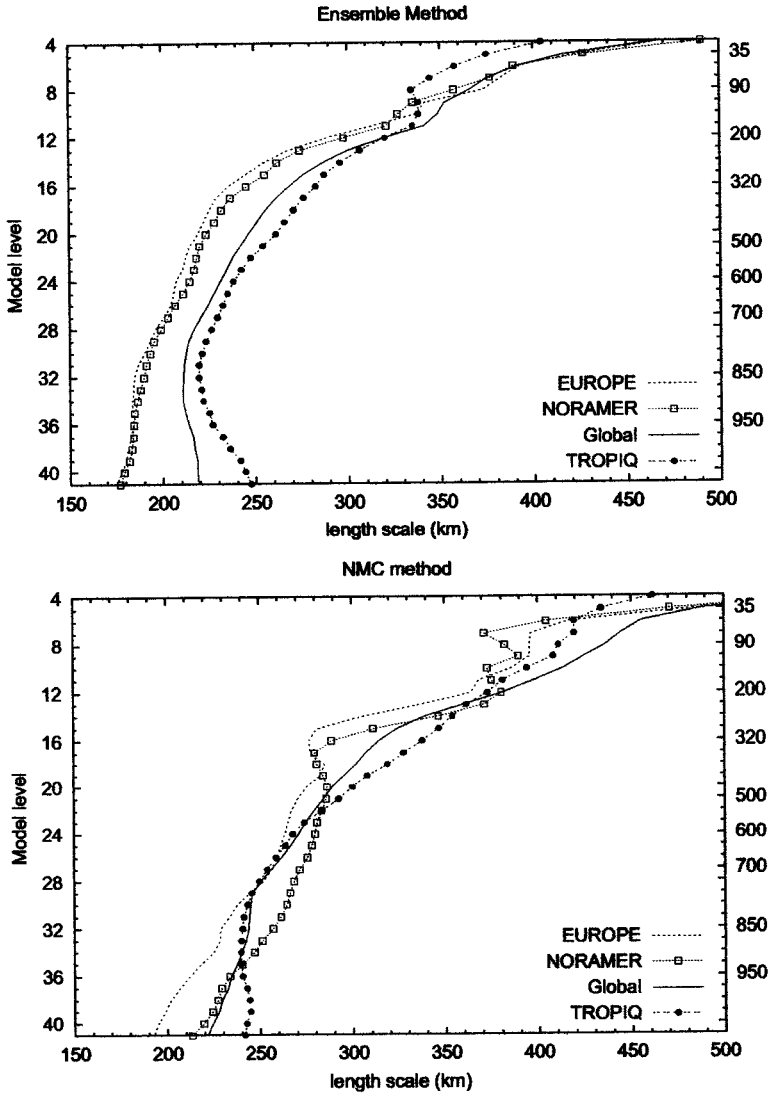


Figure 3.14: Length scales of the background error of temperature, estimated by the ensemble Method (top) and by the NMC method (bottom). The acronyms of the 3 areas are described in table 3.1.

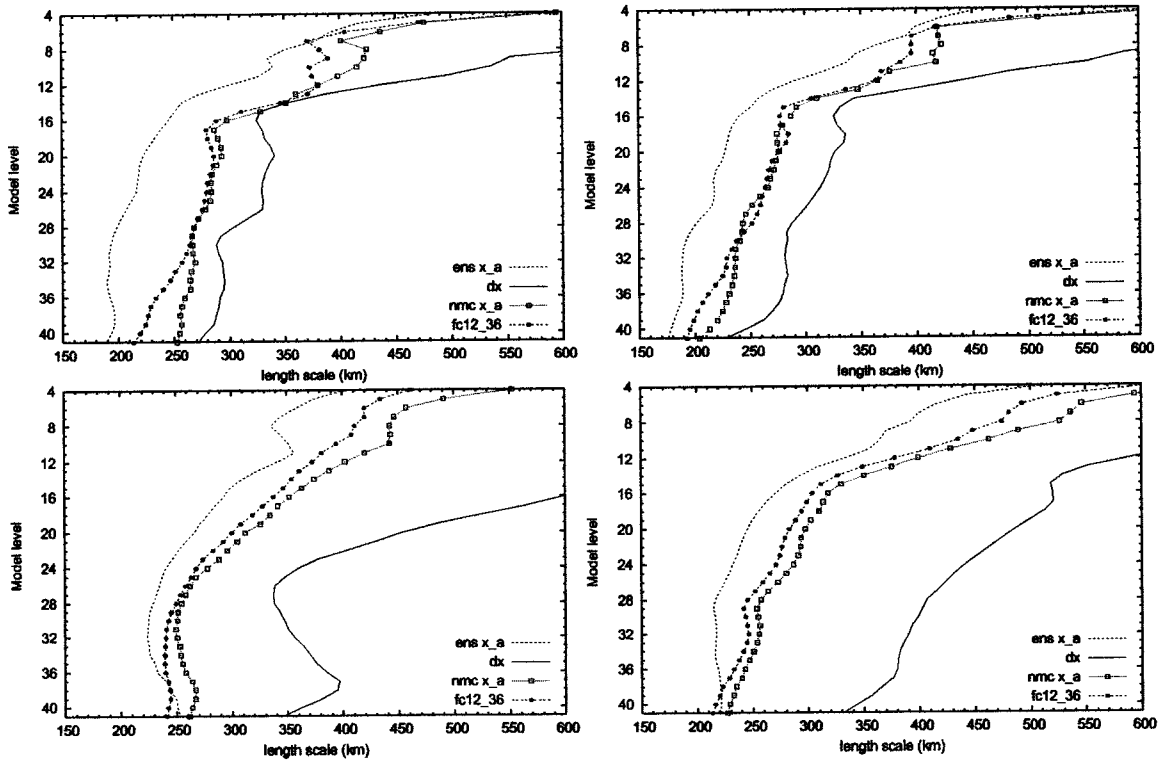


Figure 3.15: Comparison between the $L^e_{ENS}$ of the analysis error and length scales of $\tilde{\epsilon}_a^{i+4}$, $\tilde{\epsilon}_b^{i+6}$ (estimated by the NMC method) and of the analysis increments, for temperature, in: NORAMER (top left), EUROPE (top right), TROPIQ (bottom left) and SUD20 (bottom rigth).

than those from the analysis errors (see section 2.4).

From figure 3.15 it is also clear that the accumulation and time evolution of several analysis increments during a 24h period leads to a decrease of the length scales of $\tilde{\epsilon}_a^{i+4}$ when compared with those of the analysis increments, mostly in data poor areas. Nevertheless, the length scales of the analysis dispersion are smaller than those of $\tilde{\epsilon}_a^{i+4}$ for all regions. This difference is maximum over the NORAMER domain.

Furthermore, it is visible that in middle and upper troposphere the $L^e_{NMC}$ of $\tilde{\epsilon}_a^{i+4}$ and $\tilde{\epsilon}_b^{i+6}$ are very similar, specially in data rich areas.

These results suggest that the differences between the NMC and the ensemble estimates presented in figure 3.14 are mainly caused by an inadequate representation of the analysis step (which relies on the accumulation and time evolution of several analysis increments during a 24h period) in the NMC method.

### 3.4.3 Anisotropy

Hollingsworth (1987) and Bouttier (1993) have studied the anisotropy of correlation, by computing the inertia matrix of the correlation function. In a similar way, the anisotropy of $\rho$ can be determined from the eigenvectors and eigenvalues of the following matrix $\mathcal{N}$:

$$\begin{pmatrix} 1/\mathcal{N}_{xx} & 1/\mathcal{N}_{xy} \\ 1/\mathcal{N}_{xy} & 1/\mathcal{N}_{yy} \end{pmatrix} \tag{3.2}$$

where $\mathcal{N}_{xx} = L_x^2$ and $\mathcal{N}_{yy} = L_y^2$ (where $L_x, L_y$ are the zonal and meridional length scales, that are defined in section 3.4.2). The term of tilting of $\rho$ is determined by:

$$\mathcal{N}_{xy} = \frac{\sigma^2(e)}{\left\langle \frac{\partial e}{\partial x}\frac{\partial e}{\partial y} \right\rangle - \frac{\partial \sigma(e)}{\partial x}\frac{\partial \sigma(e)}{\partial y}} \tag{3.3}$$

for any error $e$, where $\sigma^2(e)$ is the variance of the background error and $\left\langle \frac{\partial e}{\partial x}\frac{\partial e}{\partial y} \right\rangle$ is the covariance between $\frac{\partial e}{\partial y}$ and $\frac{\partial e}{\partial x}$.

The magnitude of the anisotropy is measured by the oblateness ($\mathcal{O}$) of the correlation function, which is defined as $\mathcal{O} = 1 - \lambda_2/\lambda_1$, where $\lambda_1$ and $\lambda_2$ are the largest and the smallest eigenvalues, respectively. The anisotropy of $\rho$ is defined by a vector, which has the direction of the eigenvector of $\mathcal{N}$ corresponding to its smallest eigenvalue. In other words, the main anisotropy axis identifies the direction of the largest elongation of the correlation function. Moreover, the norm of the anisotropy vector is given by $\mathcal{O}$: therefore, a null vector is equivalent to an isotropic $\rho$. An oblateness equal to 0.5 means that the value of $\rho$ decreases with distance twice faster in one direction than in the perpendicular one. Thus, it is possible to diagnose the main direction and intensity of the local correlation anisotropies (see two simple examples in appendix C.2).

The figure 3.16 illustrates the kind of features that can be highlighted. For instance, the correlations appear to be SE-NW elongated over the Subtropical South Atlantic. This tilt is due to the negative (positive) covariance between the background error of the meridional and zonal components of the temperature gradient, in Northern Hemisphere (in Southern Hemisphere).

It is also interesting to mention the large anisotropy found in the region of the jet stream, particularly over North Africa and Southeast Asie (in the $10 - 20°N$ latitude belt). The zonal elongation of $\rho$ (with a SW-NE or a WSW-ENE tilt) in this area can be explained by the larger errors of the meridional component of the temperature gradient relatively to the zonal ones (see the example of $\mathcal{N}_t$ in appendix C.2). This might reflect the large uncertainties in the strength and position of the area of maximum meridional gradient of temperature associated to the African-Asian jet.

Moreover, the comparison between the correlation function at selected points (in north Africa and in Southern Atlantic) and the direction of main axis of anisotropy shows that the shape of $\rho$ is well represented by the anisotropy vector.

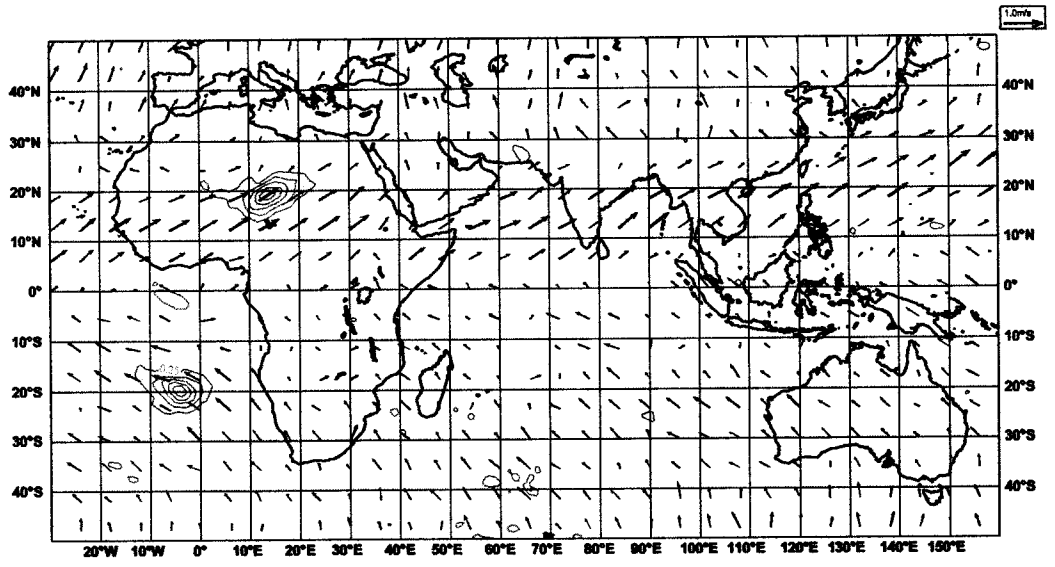


Figure 3.16: Anisotropy vector of the local autocorrelation functions of the temperature background error at model level 21, overlapped with its autocorrelation function at (20°S, 5°W) and (20°N, 14°E). The isoline interval is 0.15 starting at 0.25. The arrows are arbitrarily oriented northwards, and their scaling is such that half of the vector length from the legend corresponds to an oblateness of 0.5.

The results also show that in middle and upper troposphere, the correlation function of $u$ background error is predominantly elongated in East-West direction. This is illustrated in the top panel of figure 3.17 at model level 21. This reflects the fact that for zonal wind, uncertainties in the meridional gradient are larger than in the zonal gradient. This is consistent with the mainly zonal flow (at these levels). Moreover, this result is in agreement with other studies (Daley, 1985; Undén, 1989), which document that $\rho$ for zonal wind is predominantly elongated zonally when the wind is mainly rotational, which is the case in troposphere above the Planetary Boundary Layer.

Nevertheless, $\rho$ has a slight SW-NE tilt in Northern Hemisphere and a slight NW-SE tilt in Southern Hemisphere. Probably, this is related to the coupling between divergent and rotational components of wind (see e.g. Gustafsson *et al.*, 2001).

Near the surface, where the divergent wind component becomes more important, the shape of $\rho$ is very different, mainly over continents. For instance, the direction of main anisotropy axes is mainly Northwest-Southeast in the Rockies mountains, and mostly North-South in Andes region (not shown).

The results also show that the autocorrelation of background error for the meridional wind is mainly elongated in North-South direction, in middle and high latitudes, in middle and upper troposphere. This result is due to the fact that for meridional wind, the uncertainties in the meridional gradient are smaller than in the zonal gradient. This is consistent with the presence of transient perturbations, of the mainly zonal flow, associated to the baroclinic instabilities. Furthermore, this result is in agreement with the study of Undén (1989), which reports that $\rho$ for the meridional wind is elongated meridionally when the wind is mainly rotational, which is the case in free atmosphere in extra-tropics.

The local anisotropy diagnosis shown in this section is likely to be particularly interesting in the future, to evaluate the properties of heterogeneous covariance formulations. For instance, these diagnostic equations appeared to be very informative when studying the properties of wavelets for the error covariances of the Aladin model (Deckmyn and Berre 2005).

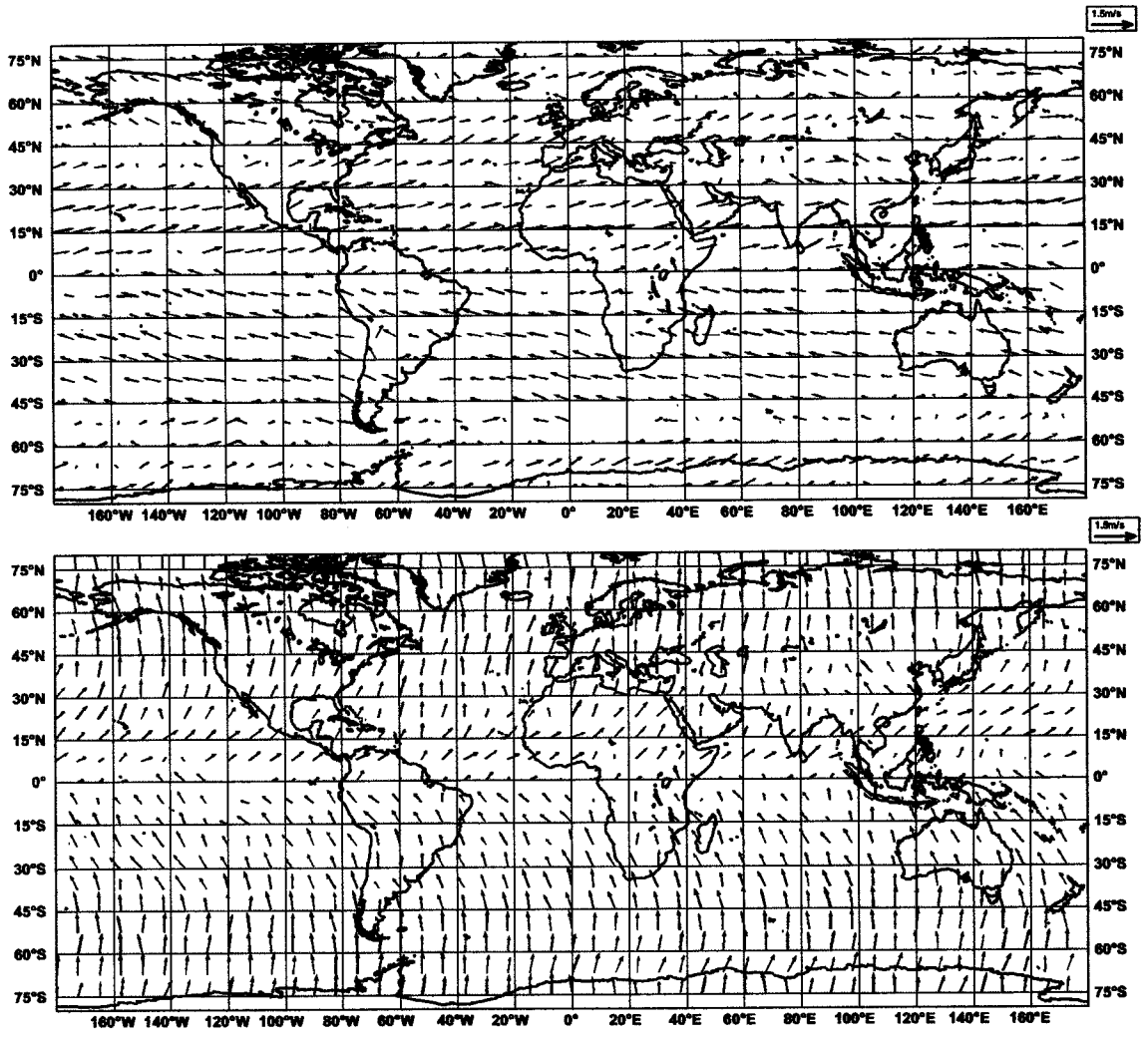


Figure 3.17: Anisotropy vector of the autocorrelation function of the background error of zonal wind (top panel) and of meridional wind (bottom panel) at model level 21 (near 500hPa).

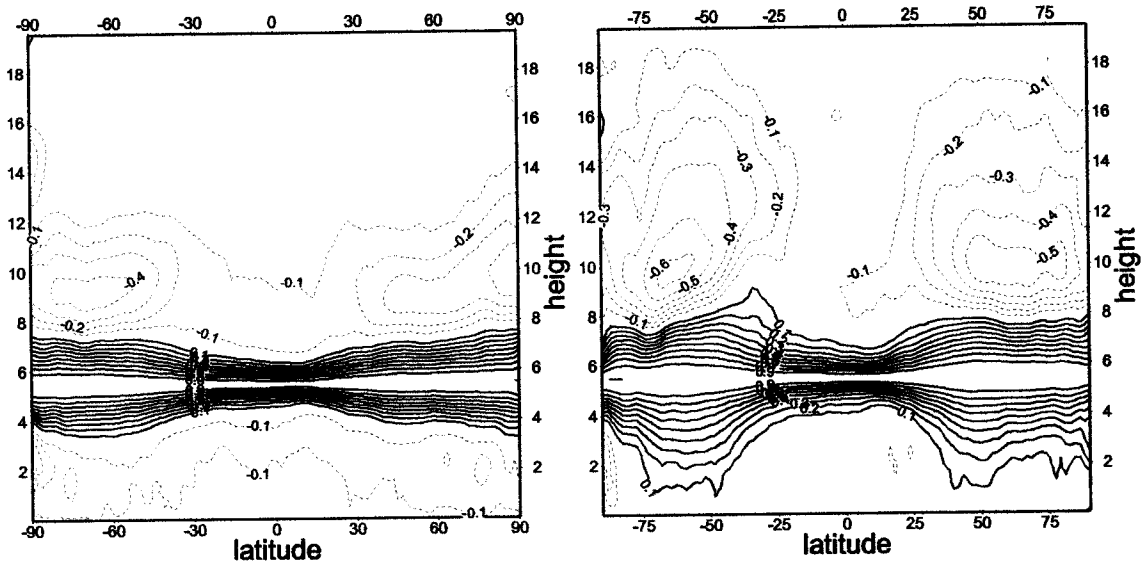


Figure 3.18: North-South variation of the vertical correlations of temperature background error (at level 21, i.e. around 500 hPa), estimated by the ensemble method (left) and by the NMC method (right). Isoline spacing is 0.1.

### 3.4.4 Vertical correlations

The latitudinal dependence of the vertical correlations of temperature is illustrated in figure 3.18. Both the ensemble and NMC methods indicate that the vertical correlations become narrower in tropics.

Moreover, it appears that the mid-tropospheric background errors are negatively correlated with levels near the tropopause. However, when the NMC method is used, the vertical correlation functions are broader, and the negative correlations near the tropopause have a larger amplitude. Furthermore, when using the ensemble method, a negative correlation between the middle and the lower troposphere is visible, while in the NMC method, the positive correlation extends from the mid-troposphere to the surface.

It is known that the vertical correlations tend to be sharper for smaller scales (for higher wave numbers), except at very large scales (Derber and Bouttier, 1999). Therefore, the fact that the vertical correlations are sharper when estimated by the ensemble method (than by the NMC method) is consistent with the enhancement of mesoscale contributions in the ensemble method.

The top panel of figure 3.19 shows the vertical correlation functions of vorticity at level 21, estimated by the ensemble method, for the global average and for the EUROPE and TROPIQ areas. It is obvious that the vertical correlations are narrower in the TROPIQ area, as expected from other studies (Ingleby, 2001; Wu *et al.*, 2002). Compared with the global average, the vertical correlation function over Europe is slightly broader in the short distances, and sharper in the long distances.

For vorticity (bottom panel of figure 3.19) and for $q$ (not shown), the background errors in the mid-troposphere are positively correlated with all levels, according to both methods. The results also show that the vertical correlations estimated by the ensemble method are narrower (than when estimated by the NMC method), for all variables.

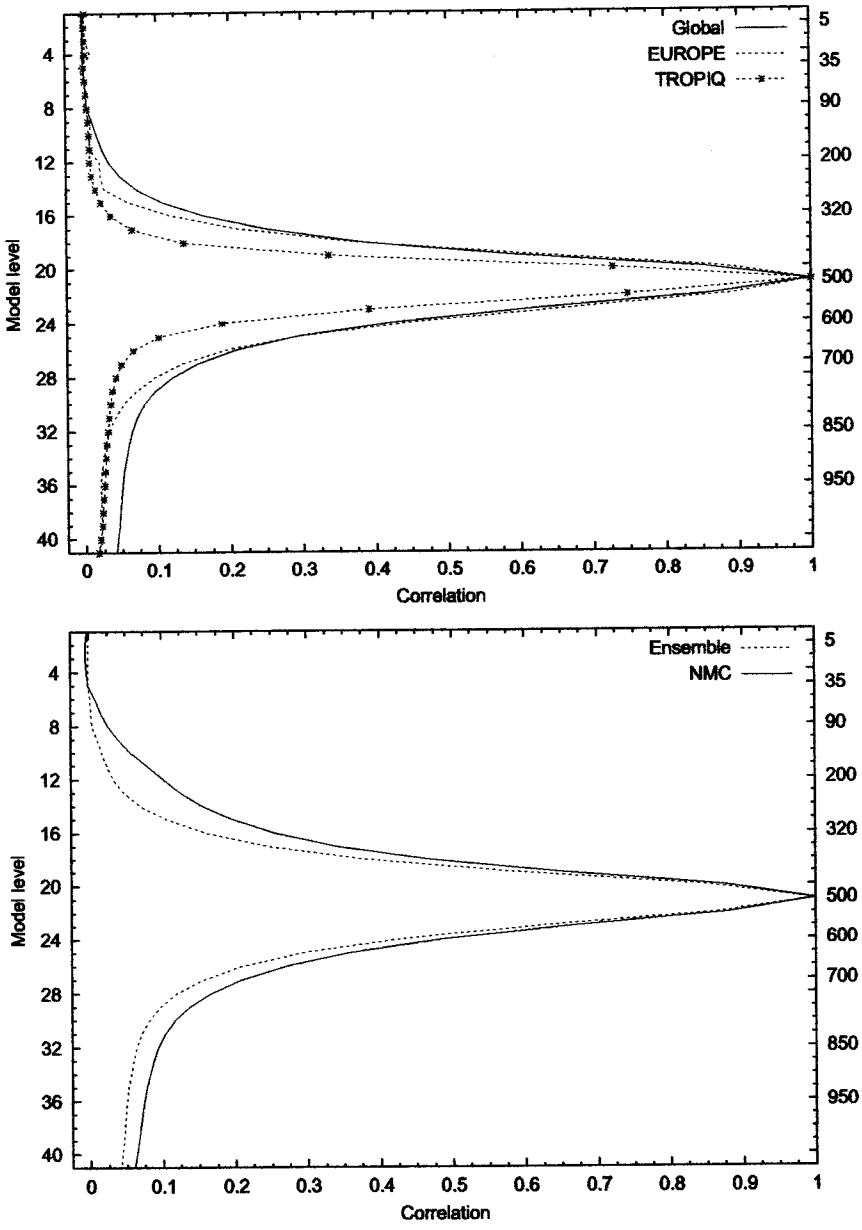


Figure 3.19: Vertical correlations of the vorticity background error (at level 21, i.e. around 500 hPa). Top panel: ensemble estimations over different parts of the globe. Bottom panel: global estimates by the ensemble method and by the NMC method.

### 3.4.5 Analysis versus background errors

The NMC and ensemble methods differ regarding the representation of the analysis step, as mentioned in section 2.4. In this section, the corresponding estimates of the local standard-deviations are compared, for both analysis errors and background errors.

In the NMC method, the vectors $\tilde{\epsilon}_a^{i+4}$, $\tilde{\epsilon}_b^{i+6}$ may be seen as estimates of the analysis errors and forecast errors, respectively (see section 2.4 : $\tilde{\epsilon}_a^{i+4}$ corresponds to some 24h-00h forecast differences and $\tilde{\epsilon}_b^{i+6}$ are the final 36h-12h forecast differences). The corresponding vertical profiles of standard deviation are represented in figure 3.20, for the SUD20 region (top panel) and for NORAMER region (bottom panel). These profiles can be also compared with the corresponding ensemble estimates, that are represented in figure 3.21. These ensemble estimates correspond to the respective dispersions of the analyses and of the backgrounds of the ensemble.

For the NMC method, it appears that the maximum magnitude of both the analysis and the background errors is slightly larger over NORAMER than in SUD20 region. On the contrary, for the ensemble method (figure 3.21), the analysis and background dispersions are larger in SUD20 than in NORAMER region. The ensemble results appear therefore to be more consistent with the expected effects of the data density contrasts (i.e. more observations over North America than in the Southern Hemisphere).

For the ensemble method, it can be also noticed, e.g. around the jet level, that the difference between the analysis and background dispersions is relatively larger over North America than in the Southern Hemisphere: as expected, the analysis effect is stronger in the data rich area than in the data poor area. Similarly, the background dispersion appears to be smaller at 18 UTC than at 12 UTC. This seems to be consistent with the larger amount of available observations at 12 UTC than at 06 UTC.

These results support therefore the idea that the analysis effects (on the error evolution) are more adequately represented in the ensemble method (than in the NMC method).

## 3.5 Impact on the forecast scores

### 3.5.1 Experiments

Some assimilation and forecast experiments have been performed in order to investigate the impact of the ensemble statistics, compared with the statistics that were used in the operational configuration (as described in appendix B). The impacts of the ensemble global covariances and of the ensemble local standard deviations have been both studied. In a similar way as for the operational local standard deviations, there are two different types of ensemble local standard deviations that are involved. Firstly, a 3D map of vorticity standard deviations is used in the minimization. In other words, this map is a part of the **B** matrix which is involved in the cost function, and which determines the weights of the background in the analysis solution. Secondly, some 3D maps of standard deviations (for wind, temperature, and geopotential) are used in the quality control. The "quality control maps" are obtained from the "vorticity maps" and from the ensemble global covariances, by using a randomization technique (Andersson and Fisher 1998). These two maps are both truncated at T21, in order to provide some smooth geographical variations.

The list of experiments and their features are summarized in Table 3.2. The experiments were performed for the period between the 5th of February and the 4th of March 2002. Some experiments have been also performed over an October/November 2002 period, and the main results were roughly similar.

Moreover, some experiments have been performed in order to investigate the impact of diurnal cycle on the ensemble statistics. For this purpose, different global and local background error statistics have been used at each analysis time in the 4D-Var assimilation cycle. The impact of these changes was mostly neutral.

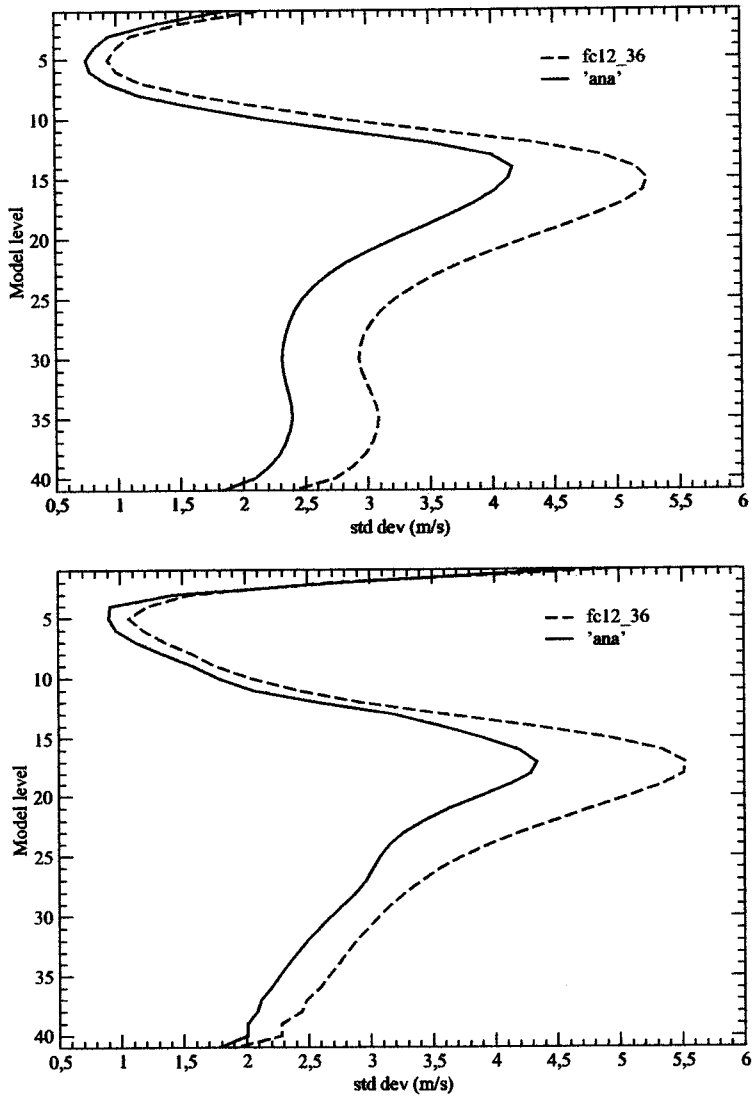


Figure 3.20: Standard deviations of the analysis error (solid line) and of the background error (dashed line), for zonal wind, estimated by the NMC method. Top panel: SUD20 region. Bottom panel: NORAMER region. The acronyms of the two regions are defined in table 3.1.

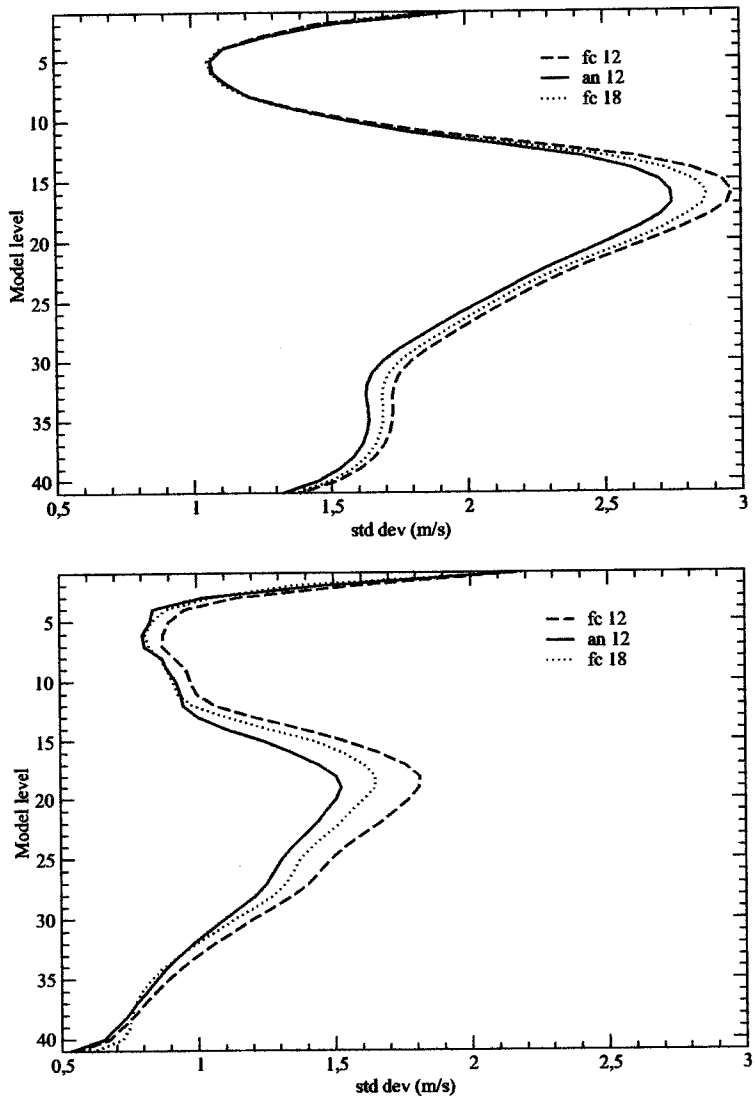


Figure 3.21: Standard deviations of the analysis error at 12 UTC (solid line), and of the background error at 12UTC (dashed line) and at 18UTC (dotted line), for zonal wind, estimated by the ensemble method. Top panel: SUD20 region. Bottom panel: NORAMER region.

In addition, a final experiment has been done in order to see the impact on forecasts of using the 3D vorticity maps provided by the NMC method. This experiment was performed for the period between the 1st and the 28th of February 2005, and compared with the operational forecasts. In operational suite at this time, the background error covariance was provided by the ensemble method. As expected, the impact of the NMC map was negative, mainly for forecast ranges larger than 48h.

Table 3.2: Experiments. The following letters are used to refer to the different statistics: ENS = ensemble statistics, NMC = NMC method's global covariances, OLS = operational local standard deviations, UNI = uniform standard deviations. The last two columns refer to the local standard deviations that are used in the minimization (min.) and in the quality control (q.c.), respectively.

| Experiment | Global covariances | Local std. dev. (min.) | Local std. dev. (q.c.) |
|---|---|---|---|
| OPE | NMC | OLS | OLS |
| EB0 | ENS | OLS | OLS |
| ER3 | ENS | ENS | OLS |
| GR3 | ENS | ENS | ENS |
| TK1 | NMC | UNI | OLS |
| ET2 | NMC | ENS | OLS |

In the experiments of table 3.2, the forecast scores were calculated by comparing the forecasts with both ECMWF analysis and TEMP observations, for geopotential, wind, temperature and relative humidity. The main results were similar with both verification sources. The impact of the ensemble statistics were mostly neutral for humidity. The largest impact in the scores was found for geopotential and wind (in a similar way). Therefore, the wind scores with respect to the ECMWF analysis will be presented, in the remainder of the text.

### 3.5.2 Impact of the ensemble global covariances

In the operational version of the Arpège 4D-Var, the global NMC standard-deviations are multiplied by a scaling factor $R_{NMC}$, which is equal to $R_{NMC} = 0.9$.

As mentioned in section 3.3.1, the ensemble estimates of the global standard deviations are much smaller than with the NMC method. A scaling factor equal to $R_{NMC} = 1.5$ has therefore been applied to the ensemble standard deviations: the resulting magnitudes are then similar to the operational ones. In addition, some sensitivity studies have indicated that this choice of $R_{NMC}$ was close to optimal (the scores were slightly degraded or neutral, when decreasing or increasing $R_{NMC}$ by 10%).

The left panels of figure 3.22 correspond to the impact of the ensemble global covariances, compared with the NMC global covariances. The impact appears to be clearly positive over the AMNORD and EURATL domains, with some increase of the impact amplitude when the forecast range increases. The scores are also generally positive over the tropics. In the other regions, the scores are slightly positive or neutral.

### 3.5.3 Impact of the ensemble local standard deviations

The middle panels of figure 3.22 correspond to the total impact of both the ensemble global covariances and the ensemble local standard deviations (of vorticity). The generally positive

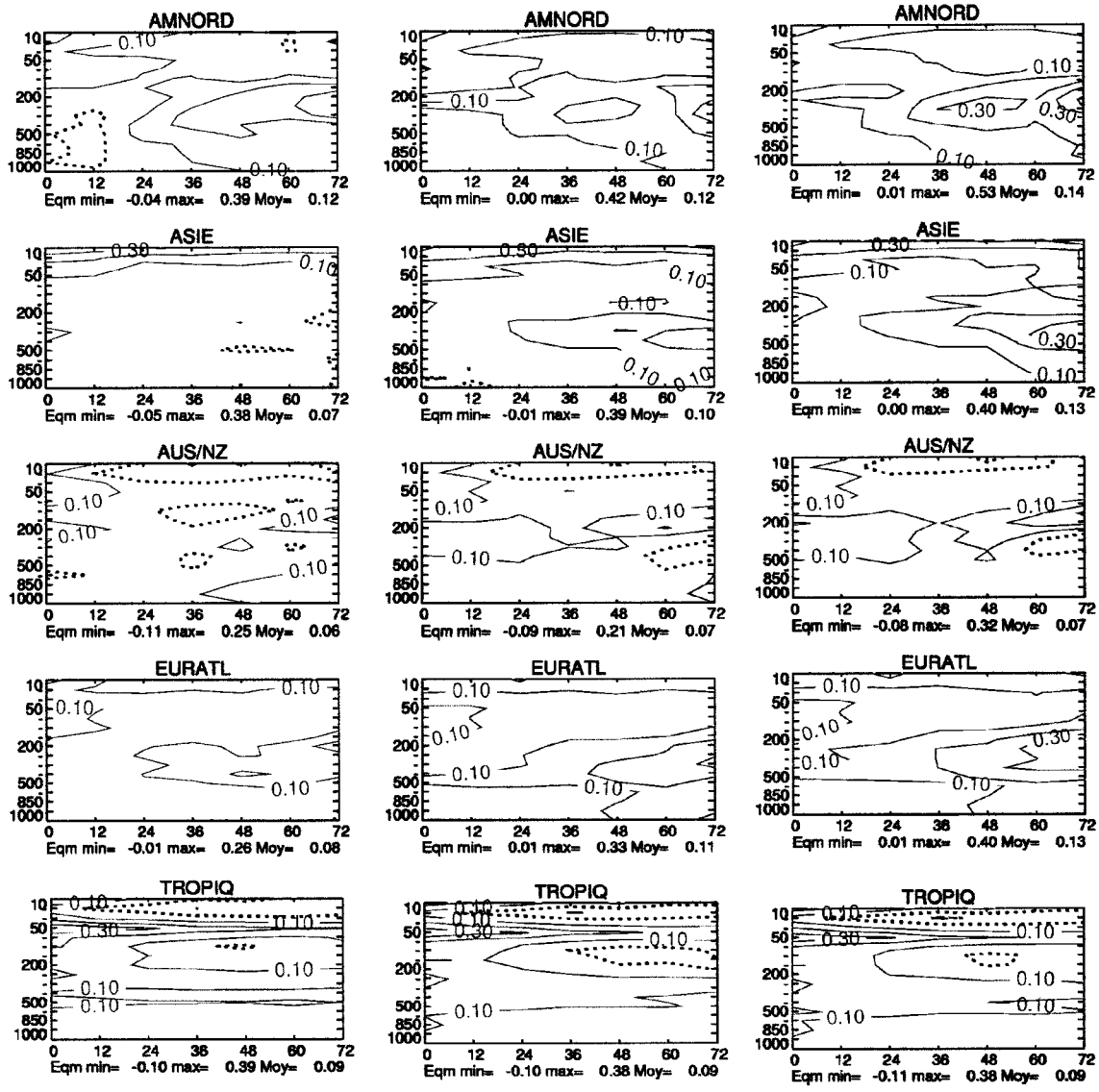


Figure 3.22: Difference between the root mean square errors (RMSE, in m $s^{-1}$) of wind forecasts (as function of pressure levels and forecast range) of the operational experiment (OPE) and those of the following experiments: EB0 (left panels), ER3 (middle panels) and GR3 (right panels). Forecasts are verified against the ECMWF analysis. Solid (resp. dotted) isolines indicate ranges and levels, for which the scores of the involved experiment are better than (resp. similar to) the scores of the operational experiment. The isoline spacing is 0.10 m $s^{-1}$. The acronyms of the 5 areas are defined in table 3.1.

impact, which was already visible in the left panels, appears to be enhanced due to the use of the ensemble local standard deviations. This is particularly noticeable over the ASIE and EURATL areas. This result suggests that the enhancement (by the ensemble method when compared with the former operational map) of the standard deviation in these regions (see figure 3.8) is relevant. A similar positive impact was noticed when comparing the experiments ET2 and OPE.

The impact of the ensemble local standard deviations, that are used in the minimization, has been also studied by comparing the experiments ET2 and TK1 (not shown): TK1 is based on some uniform standard deviations (i.e. that do not vary geographically), while ET2 is based on the ensemble local standard deviations (see table 3.2). The positive impact was confirmed, and it appeared to be even stronger than when comparing the ensemble and operational standard deviations (which is also an indication that the operational standard deviations are more realistic than the uniform standard deviations).

The right panels of figure 3.22 correspond to the total impact of the three types of ensemble statistics, compared with the operational statistics: the third involved ensemble statistics are the local ensemble standard deviations, that are used in the quality control stage of the 4D-Var analysis. The addition of this third ingredient appears to strengthen the previous positive impacts. This is more obvious over the AMNORD, ASIE and EURATL areas.

## 3.6 Summary

In the present chapter, the results of the analysis ensemble experiments performed with the non-stretched version of the Arpège global model, in a perfect-model context, have been presented.

Globally, the ensemble approach (when compared with the NMC method) appears to emphasize the relative contributions of the small scales to the background errors. Consequently, the correlation functions are sharper in the ensemble method. These results are in accordance with the ones described by Fisher (2003), where these differences were considered to be caused by the involvement of longer forecast ranges in the NMC method. Nevertheless, the results presented here suggest that the different representations of the analysis step (in the simulation of the error evolution) explain also an important part of the differences between the two methods.

The examination of the local covariances suggests that the analysis error variances estimated by the NMC method are underestimated in the data-poor areas, and overestimated in regions where the observations are dense and have a high quality. This result is consistent with the theoretical considerations made by Bouttier (1994).

In comparison, the local ensemble background error variances are able to capture interesting features, such as the contrasts between data-rich and data-sparse regions, and also the areas of large atmospheric variability (the ITCZ, for instance). Moreover, the analysis dispersion provided by the ensemble approach appears to give some relevant information about the analysis error, for instance concerning the influence of data density.

The local correlation structures were also examined. An economical method was presented to estimate the local correlation length scale. The corresponding estimates provided by the ensemble method show that the length scale is largest in the tropics and smallest in the data rich regions, as it would be expected from other studies (e.g. Lindzen and Fox-Rabinovitz (1989), Ingleby (2001), Bouttier (1994)). In addition, these estimates from the NMC method differ from the ensemble ones, mainly for temperature. In particular, the results indicate that the overestimation of the length scales by the NMC method is more pronounced over data rich areas, such as North America. These differences suggest that the expected data density effects on the correlation structures are not well represented by the NMC method.

Moreover, this local length scale diagnosis allows also to examine the correlation anisotropies. These informations about the heterogeneities and anisotropies can be useful to model spatially varying correlations in the **B** matrix, by wavelet techniques (Fisher (2003), Deckmyn and Berre

(2005)) for instance.

Some impact studies were finally performed with the Arpège 4D-Var. The use of the ensemble global covariances (compared with the NMC global covariances) had a general positive impact on the forecast quality. This positive impact was found to be enhanced when the ensemble local standard deviation maps (instead of the former operational ones) are used.

The diagnostic and impact studies indicate that there is some relevant global and local information in the ensemble method, and in particular that this technique is more appropriate than the NMC method. Nevertheless, the use of the ensemble approach in a perfect model-framework might lead to an underestimation of the error variances. In the next chapter, the impact of the model error on the background error covariances will be presented.

# Chapter 4

# The role of model error in the ensemble approach

## 4.1 Introduction

The study of the background error covariances of the Arpège global NWP model, estimated by the ensemble approach using a perfect-model assumption, has been presented in the previous chapter. The assumption that the model is perfect means that the imperfections of the forecast model are not considered to be a source of error to the background error. In others words, it is supposed that the analysis and background errors result from observational errors, from induced errors in the background field used by the data assimilation, and from their evolution during the analysis and forecast (from a perfect model) steps.

According to Houtekamer *et al.* (1996), neglecting the model error in the ensemble simulations might lead to an ensemble dispersion that would be too small. The results presented in the previous chapter also support this idea. Houtekamer *et al.* (1996) also suggest that, in order to obtain a realistic estimate of the background error covariances, it may not be necessary to have a complete description of the model error. Instead, it could be sufficient to concentrate on aspects of the model which can be expected to be deficiently represented. Moreover, one may assume that differences between two model versions characterizes the model error due to the parametrized processes.

In the present chapter, the effect of the simulated model error on the background error covariances, estimated by the ensemble method, is investigated. It is assumed that the horizontal discretization and the uncertainties associated to the physical parametrizations are significant sources of model error. The role of horizontal discretization has been studied by performing experiments with different horizontal resolutions. In addition, the impact of the uncertainties in the parametrizations has been examined by performing experiments using different values for certain tuning parameters from some parametrizations.

The description about the changes made in the chosen parameters and parametrizations is presented in appendix E. The list of experiments and their features are summarized in Table 3.2. These experiments were performed for the same period as the ones using the perfect-model assumption (1st of February to 24th of March 2002).

## 4.2 Global background error covariances

This section presents the comparison between the background error covariances estimated in the spectral space (at truncation T179) by the three types of ensembles described in table 4.1.

Table 4.1: Ensemble experiments performed to study the impact of simulating the model error. Here, HD is the horizontal diffusion and LSCON is the large scale convergence used in the parametrization of convection. The truncations T299 and T449 correspond to a grid spacing of approximately 67km and 45km, respectively.

| EXP | Member | Observations perturbations | Resolution | Changes in parametrizations | Error source |
|---|---|---|---|---|---|
| PO | P33 | yes | T299 | HD(T299) | observations |
| | P34 | yes | T299 | HD(T299) | |
| POR | A12 | yes | T449 | HD(T449) | observations |
| | P34 | yes | T299 | HD(T299) | and resolution |
| PORP | A15 | yes | T449 | HD(T449)<br>vertical diffusion<br>cloudiness scheme<br>LSCON | observations,<br>resolution<br>and<br>physical |
| | P35 | yes | T299 | HD(T299) | parametrizations |

### 4.2.1 Standard deviations and covariances

The figure 4.1 presents the vertical profile of standard deviations of temperature and vorticity background errors, estimated by ensemble method for the three types of ensemble differences described in table 4.1. There are some similarities between the three ensembles, mostly in the upper troposphere and in the stratosphere, where the variations of the standard deviation with height are very similar. Another common feature is that the background errors of vorticity increase significantly with height, reaching their maximum near the tropopause.

Nevertheless, the standard deviations are largest for PORP and smallest for PO. In other words, the background errors are largest when the model error due to resolution and parametrizations uncertainties is simulated. Moreover, the results suggest that for vorticity and divergence (not shown) the resolution plays the main role. This is consistent with the fact that for these variables the largest contributions to the background error are due to the mesoscale phenomena.

For temperature (figure 4.1, left panel) it is interesting to note that the model error simulation has not only an impact on the magnitude of the background errors, but also on its variation with altitude. For instance, according to PO, in the troposphere, the largest background errors are found in middle troposphere. To contrast with this, for PORP, in troposphere, the largest temperature background errors are located in the PBL and decrease with height. This results seems to be more realistic than those obtained for the other ensembles, when considering the great importance of turbulent fluxes in PBL and that the turbulent motions have spatial and temporal variations at scales much smaller than those represented in the analysis (due to the resolution of the meteorological observing network). This is in accordance with the larger contributions of the mesoscale phenomena to the background error in the bottom model levels (relatively to the middle and upper troposphere), as shown in figure 3.3 of the previous chapter. In addition, the results for PORP are consistent with those presented in the appendix D, where it is shown that the RMSE of the innovation vector for temperature has larger values near the surface than in middle troposphere.

The figure 4.2 presents the standard deviation of the background error of $P_s$ as a function of the wave number (or horizontal scales). One can see that the simulation of model errors contributes to increase the magnitude of surface pressure background error. This is more pro-

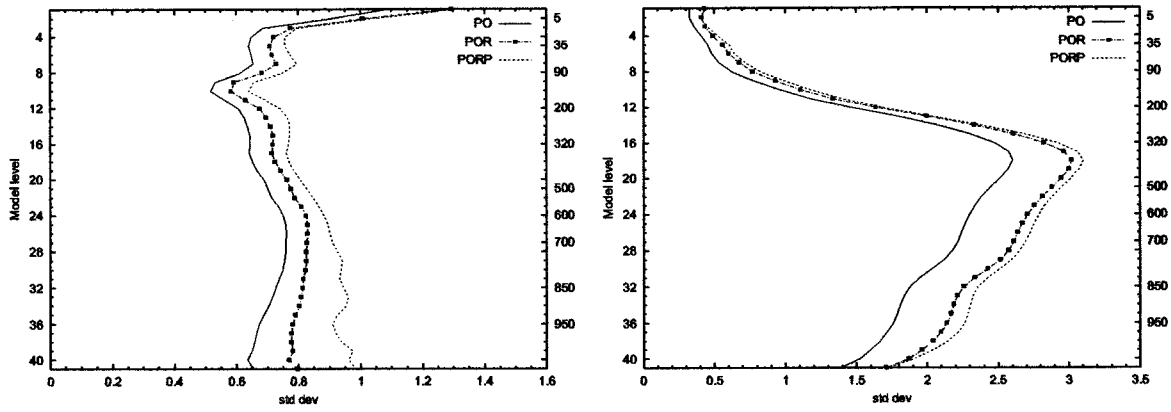


Figure 4.1: Vertical profile of standard deviation of the temperature (K) (left side) and vorticity ($10^{-5}s^{-1}$) (right side) background errors, estimated by the ensemble method for the experiments described in table 4.1.

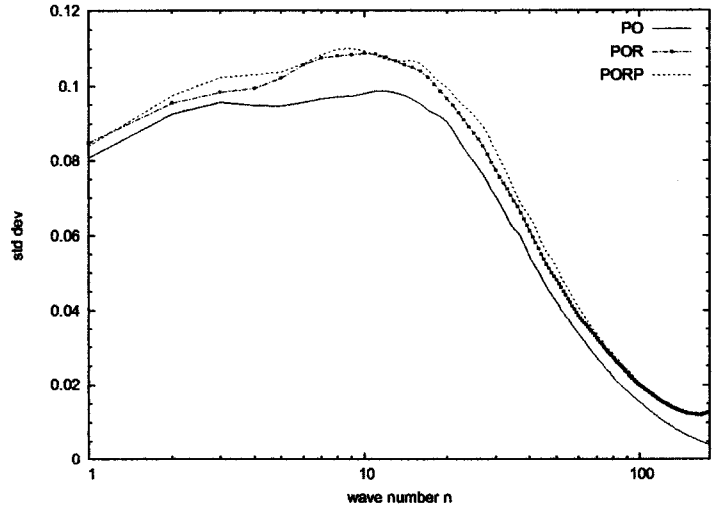


Figure 4.2: Standard deviation of surface pressure background error as function of the wave number, estimated by the Ensemble method, using the PO, POR and PORP ensembles.

nounced for synoptic scales (for wave numbers between 4 and 20) and for small scales (for wave numbers greater than 100). In section 4.3.1 it will be shown that this enhancement of the small scales contributions is associated to the increase of the background errors over mountainous regions, when the model error is simulated.

The figure 4.3 presents the variance and correlation spectra for temperature at model level 32 (near 850hPa). All ensembles agree that the largest variance of the temperature background error is caused by the synoptic scales. However, for PORP the contributions from synoptic scales are enhanced, relatively to the small scales (bottom panel). This result can be surprising, because one could expect that the role of the model error would be to emphasize the variance associated to the mesoscales. However, the fact that the errors in smaller scales tend to propagate to larger scales due to non linear terms can possibly explain this result.

As illustrated in figure 4.4, this enhancement of the mesoscale contributions implies that the corresponding correlation length scales (computed from equation 1.48, section 1.3.3, and defined by equation 1.42, section 1.3.2) are larger for PORP than for the other ensembles, in the troposphere.

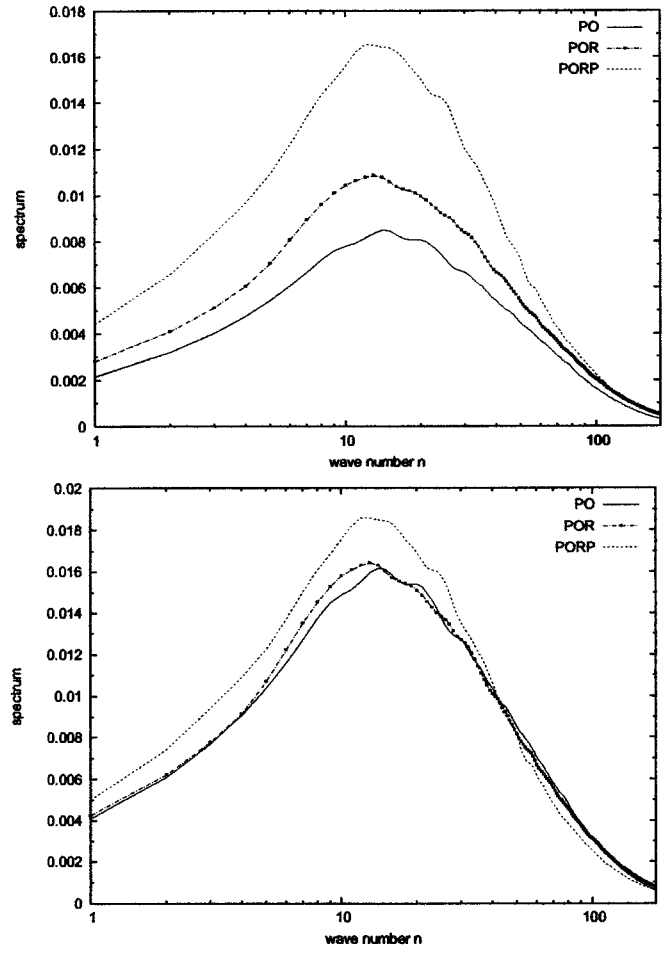


Figure 4.3: Variance spectra (top) and correlation spectra (bottom) of temperature background error at level 32 (near 850hPa), estimated by the ensemble method for the experiments described in table 4.1.

Moreover, the larger difference between PORP and POR (than between POR and PO), illustrated in figures 4.1 (left panel) and 4.3, suggests that in lower troposphere, the model error caused by uncertainties in the physical parametrizations has a larger contribution to the temperature background errors than the model error related to resolution. Considering that relatively high resolutions have been used in the experiments, and that the largest variance of temperature errors is due to the synoptic scales, one could expect that, if coarser resolutions had been used, the model error sensitivity to resolution would have been larger.

The results also show that the simulation of model errors in the ensemble method contributes to increase the magnitude of specific humidity background error (not shown).

The figure 4.5 presents the correlation spectra for vorticity at model level 32 (near 850hPa). One can see that compared with the PO, ensembles POR and PORP emphasize the relative contributions of the small scales (compared with the contributions of the large scales). The same was found for divergence. Thus, the role of model error is to enhance the relative contributions of small scales to the background errors of vorticity and divergence, mostly due to the model error associated to the horizontal discretization. As a consequence, the vorticity and divergence length scales are smaller for the ensembles for which the model error is simulated than for PO (not shown). Another consequence is that the contributions of the unbalanced components are emphasized in the multivariate formulation (see figure 4.6). This is particularly marked in

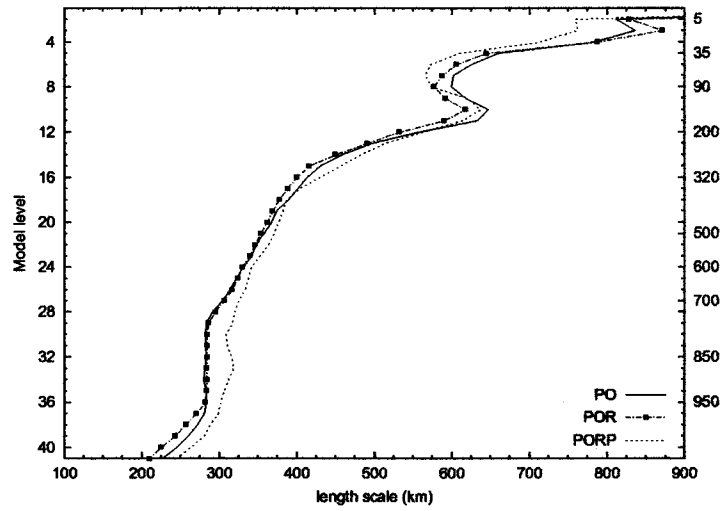


Figure 4.4: Length scale of temperature background error, estimated by the ensemble method for the experiments described in table 4.1.

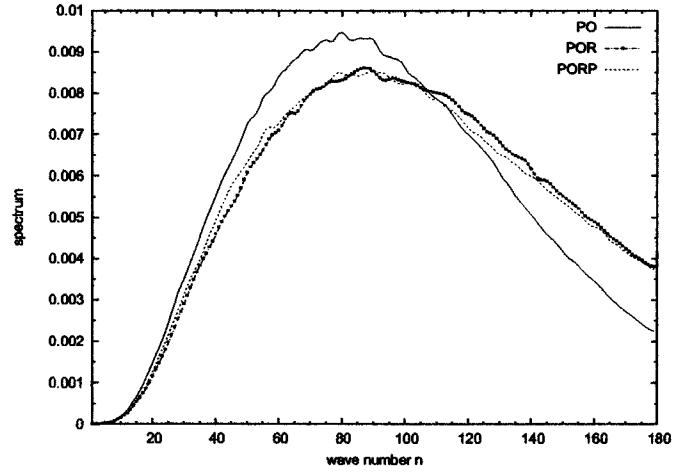


Figure 4.5: Correlation spectra of vorticity background error at level 32 (near 850hPa), estimated by the ensemble method for the experiments described in table 4.1.

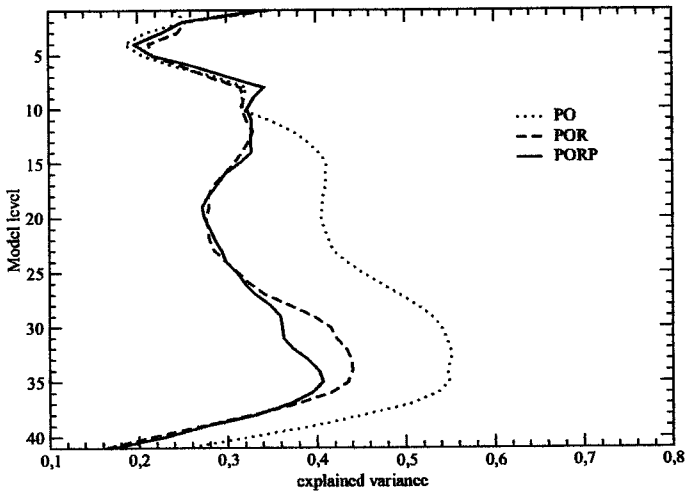


Figure 4.6: Explained variance for temperature background error, for the experiments described in table 4.1.

mesoscales, for wave numbers larger than 60 (not shown). Moreover, this result is mostly due to the contributions of model error caused by resolution.

### 4.2.2 Analysis versus background error

As said in the previous chapter, the differences, $\varepsilon_a$, between different perturbed analyses may be seen as an estimate of the analysis error ($e_a$). In other words, the dispersion of the analyses of the ensemble corresponds to an estimate of $e_a$. Thus, the comparison between $\varepsilon_a$ and $\varepsilon_b$ gives the opportunity to compare the simulated analysis and background errors and to study the impact of the analysis step in the ensemble simulations.

The corresponding vertical profiles of the standard deviation for temperature are represented in figure 4.7 for PO and PORP experiments. According to both ensembles, in the upper and middle troposphere, the analysis error is smaller than the background error. However, for PO and for POR (not shown), in the PBL, the analysis dispersion is equal or slightly larger than the background dispersion. This result is unforeseen, because one would expect that the analysis process will reduce the background uncertainties thanks to the information from the observations (see sections 1.1 and 1.2). To contrast with this, for PORP, the analysis dispersion is smaller than the background dispersion also in the PBL. This result is also illustrated in the bottom panel of figure 4.8 for the model level 32 (near 850hPa).

One can see that, for PORP, the analysis dispersion is smaller than the background dispersion for all different scales. The largest difference between the simulated errors is found for synoptic scales. On the contrary, for PO and POR ensembles, the analysis error is larger than the background error, mostly for synoptic scales. This may be explained as follows: for PORP the synoptic scale contributions to the background errors are largely emphasized, while the mesoscales become less important; and it is expected that the analysis system will reduce the model state uncertainties more efficiently for synoptic scales than for mesoscales, partly due to the density of the observation network, and partly due to the filtering properties of the analysis (see section 1.3.4).

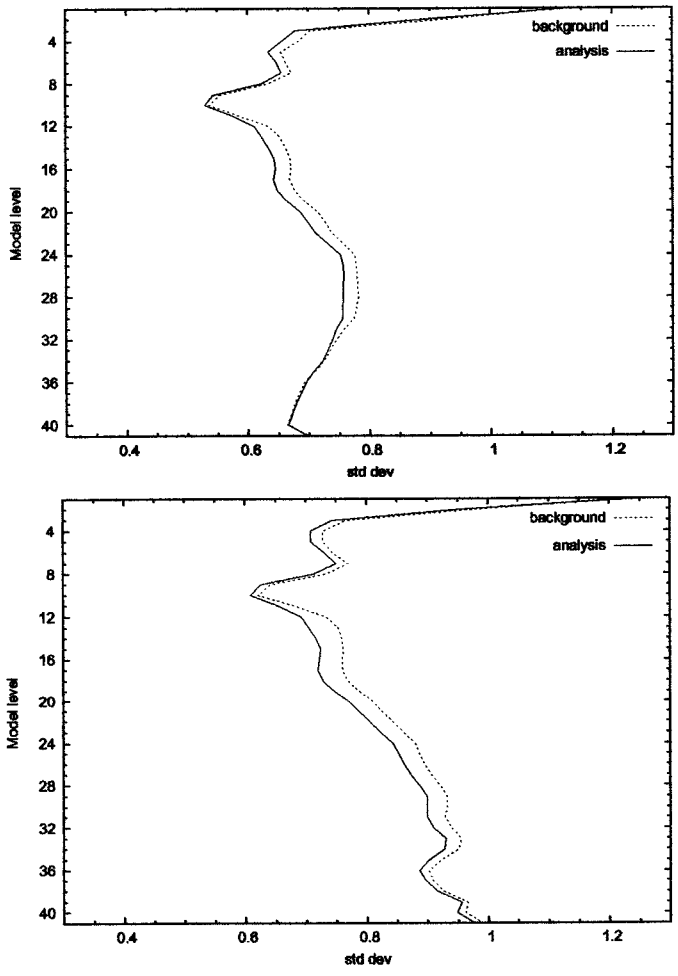


Figure 4.7: Standard deviations of the analysis error (solid line) and of the background error at 12UTC (dashed line), for temperature, estimated by the ensemble method with PO (top panel) and PORP (bottom panel) ensembles.

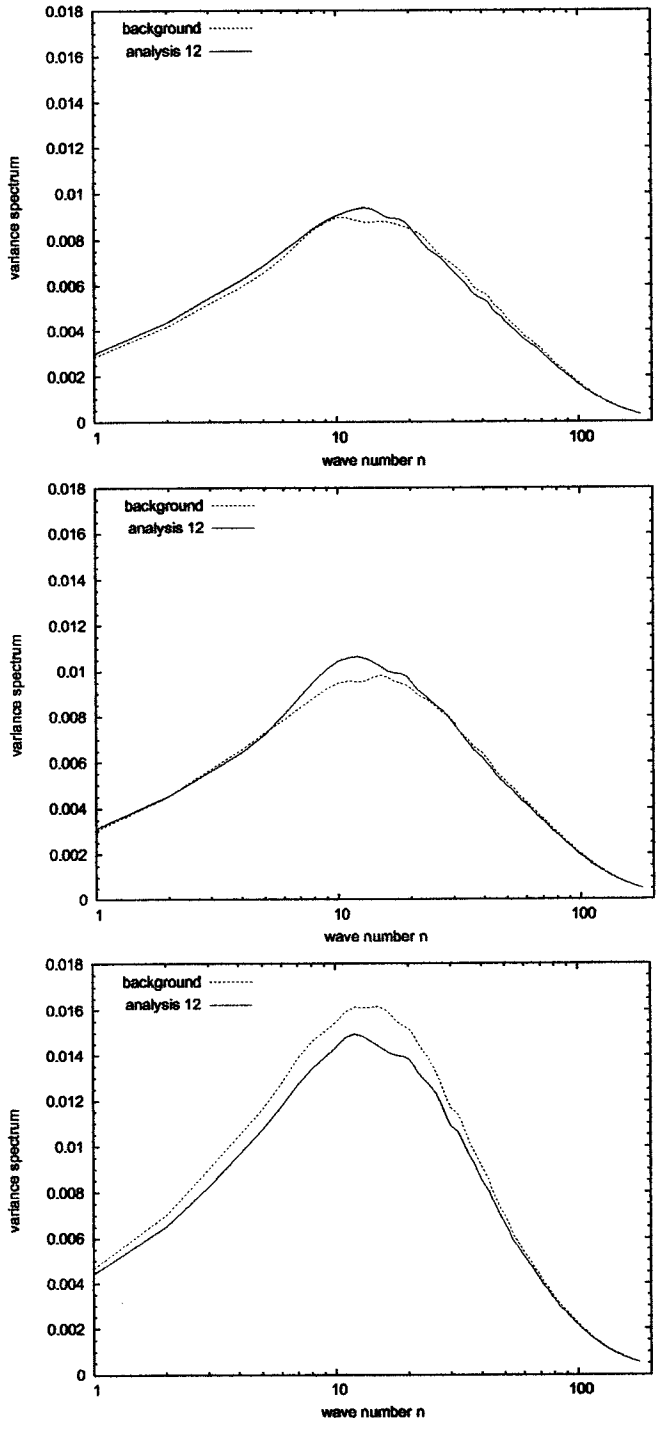


Figure 4.8: Spectra of analysis error (solid line) and of background error valid at 12UTC (dashed line), for temperature at level 32 (near 850hPa), estimated by the ensemble method with PO (top), POR (middle) and PORP (bottom) ensembles.

## 4.3 Local background error covariances

### 4.3.1 Standard deviations

The maps of standard deviation of the surface pressure background error for the three ensembles described in table 4.1 are shown in figure 4.9. Surface pressure is a large scale variable, therefore it was decided not to filter its background error standard deviations.

Some common features (between the 3 ensembles) can be identified regarding the latitudinal variations and the influence of the data density: the standard deviations are smaller in the tropics than in middle and high latitudes, and they are also relatively small over data rich regions, such as Europe.

However, when the model error is simulated, the standard deviation is larger over land, mostly in mountainous areas. This feature seems to be more realistic, since the atmospheric evolution is more difficult to predict in such regions. Moreover, the local maxima over the mid-latitude oceans are enhanced, such as over the Northern West Pacific and Northern Atlantic.

The figure 4.10 presents the vertical profile of the standard deviations of vorticity background errors, for the three ensembles, for NORD20, TROPIQ and SUD20 regions. All ensembles agree that the largest variance occurs in SUD20 and the smallest in TROPIQ region. This is consistent with the largest variability of the vorticity field in middle and high latitudes (associated to frontal systems, extra-tropical cyclones and jet streams). This result is also consistent with the smaller density of observations in the Southern Hemisphere than in the Northern Hemisphere.

The results also show that the enhancement of the magnitude of the background error due to the simulation of model error is stronger in NORD20 than in SUD20 domain. This is related to the larger percentage of land over NORD20, where the role of model error seems to be more important. In these two domains, the model error related to resolution appears to be more important than the one associated to parametrizations. However, in TROPIQ region both factors are relevant.

The normalized standard deviations of the vorticity background error were also examined for the three ensembles. The corresponding maps for level 21 (near 500hPa) are shown in figure 4.11. Some similarities are visible concerning the latitudinal variations and the data density contrasts: the standard deviations are smaller in the tropics than in the middle latitudes, and they are also relatively smaller over data rich areas, such as Europe and Northern America.

On the other hand, for the PORP ensemble, the local minima are slightly more marked over the tropics, and slightly less pronounced over the data rich areas. In addition, the local maxima over Northern Atlantic and over Himalayas are enhanced. The local maxima near the North Pole are less pronounced. For the POR ensemble, the standard deviation map is very similar to the one from the PORP ensemble, except that for the former the local maximum over Himalayas is stronger.

The figure 4.12 presents the vertical profile of the standard deviations of temperature background error, estimated by the ensemble method for the three ensembles, for different regions (see table 3.1, section 3.4.2). In SUD20 region (top left panel), the shape of the vertical profile of standard deviation is very similar for PO and POR ensembles. For instance, the smallest magnitude of the background errors is found near the surface. Nevertheless, the values of the standard deviations are larger for the POR set. The simulation of the model error due to uncertainties in physical parametrizations causes a further increase of the standard deviation values, mostly near the surface (mainly over land regions, not shown), which contributes to change the shape of the vertical profile.

In TROPIQ region, all ensemble simulations agree that large errors are located near 90hPa (top right panel). Near the equator, the tropopause is located approximately at 100hPa. Thus, the large errors near 90hPa probably reflect the difficulties of the model in determining the

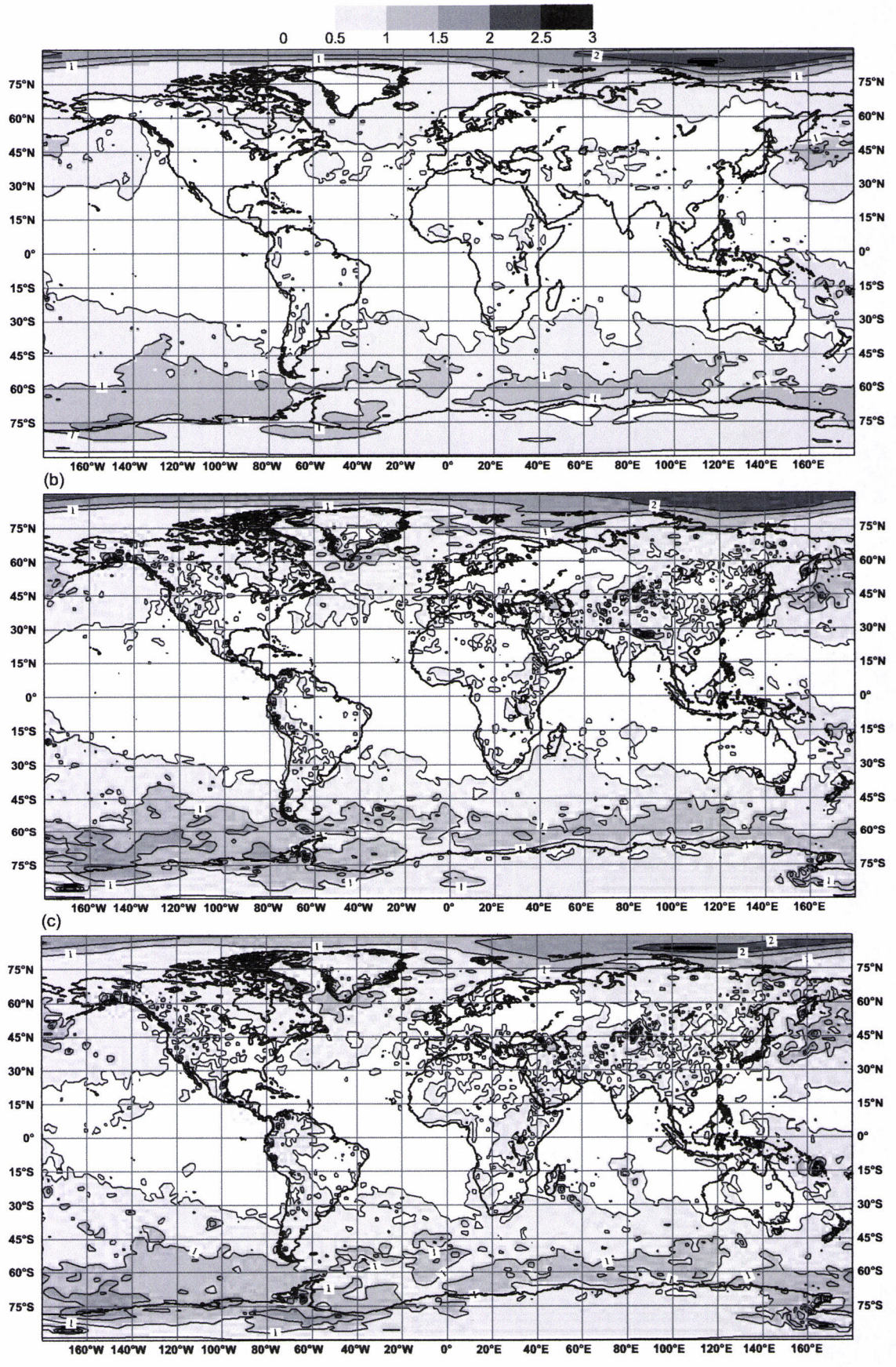


Figure 4.9: Standard deviations of the surface pressure background error (hPa), for (a) PO, (b) POR and (c) PORP ensembles.

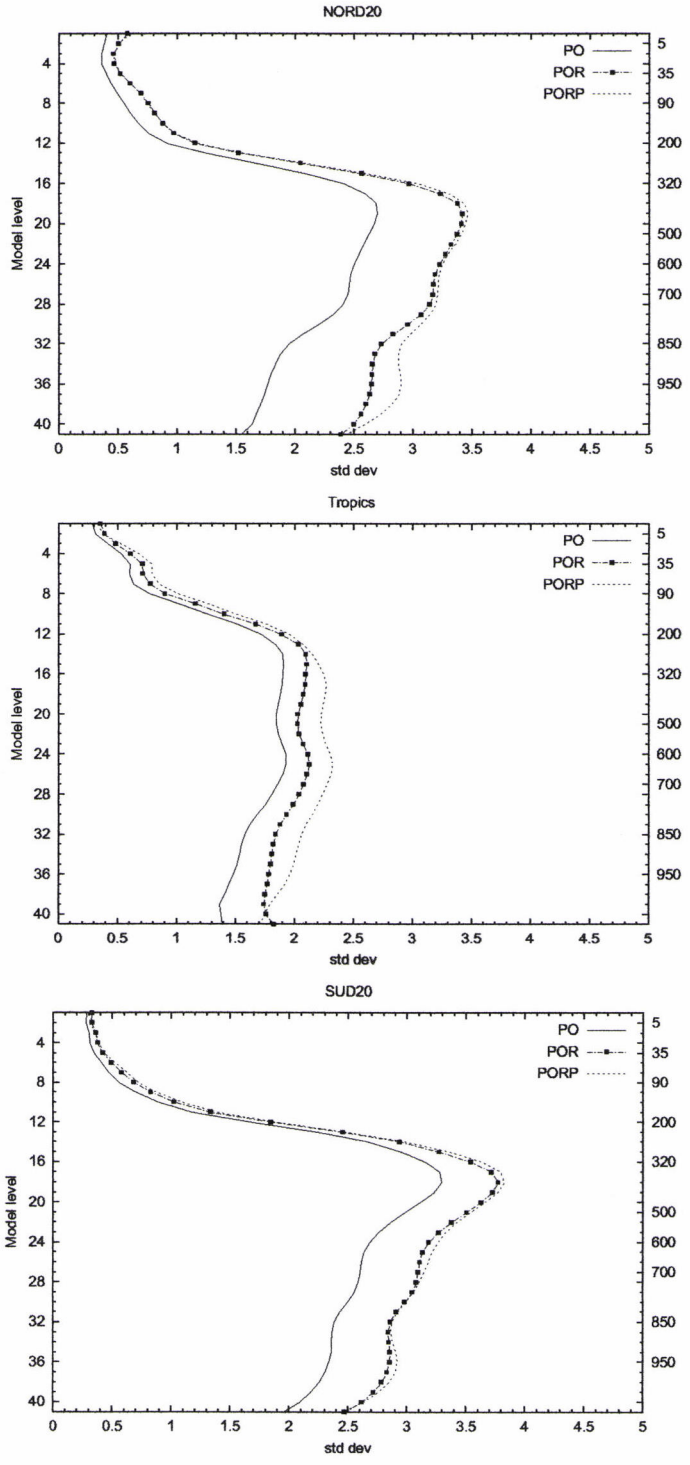


Figure 4.10: Vertical profile of standard deviation of vorticity ($10^5 s^{-1}$) background error estimated by Ensemble methods.

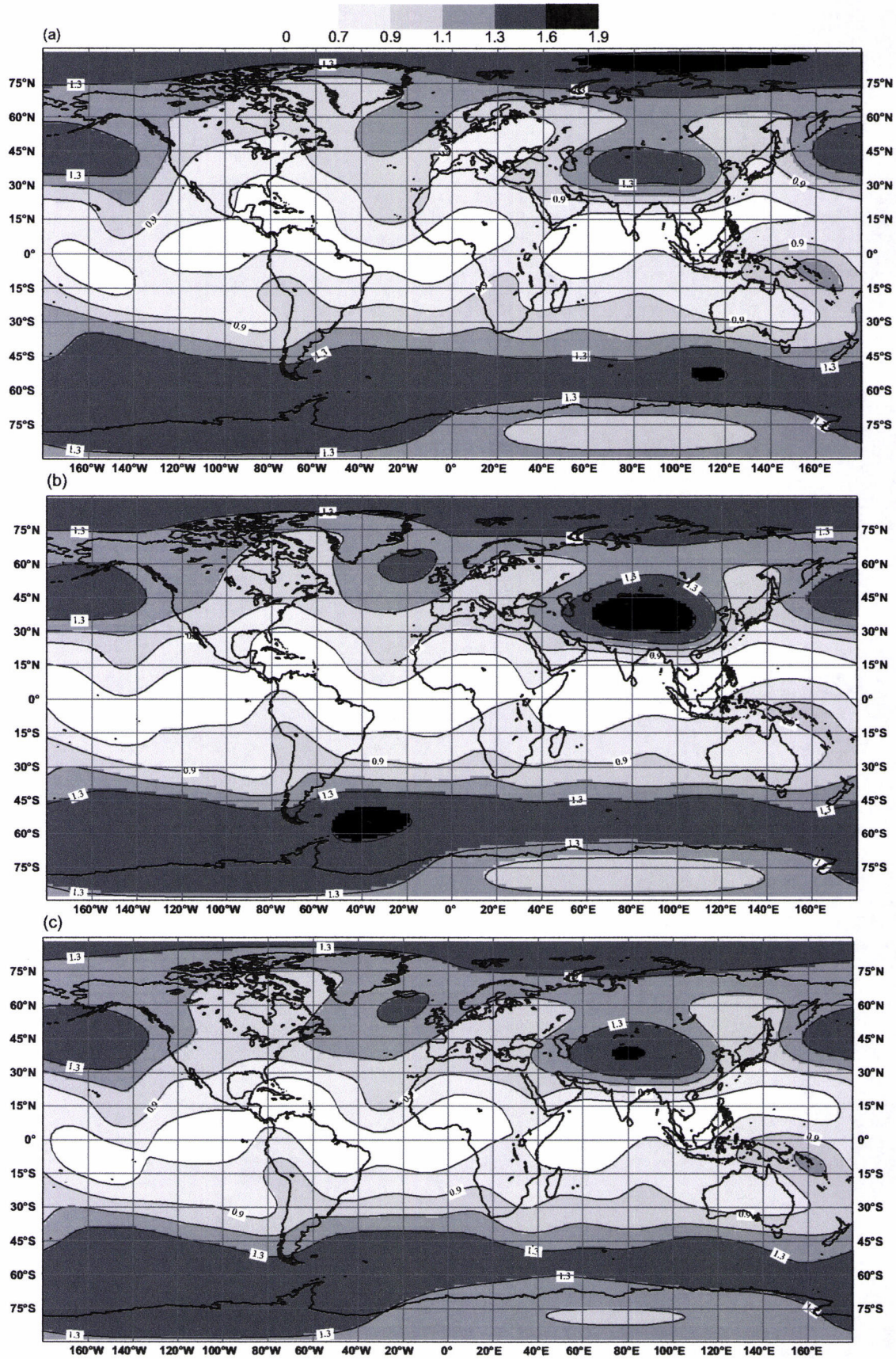


Figure 4.11: Normalized standard deviations of the vorticity background error at model level 21 (around 500 hPa), for (a) PO, (b) POR and (c) PORP ensembles. (These maps were produced with a spectral truncation equal to T21.)

position of tropopause. Possibly, the fact that the ARPEGE model has a lower vertical resolution at these altitudes contributes to this result.

Moreover, in the tropics, it is visible that, in upper troposphere, the simulation of the model error due to uncertainties in physical parametrizations causes a considerable increase of standard deviations. This result is consistent with the fact that the vertical gradients are strongest in the tropical upper troposphere, due to the subsidence in the neighbouring of the deep cumulonimbus clouds in the ITCZ, and due to the evaporation and radiative cooling at the tops of these clouds.

In troposphere, over the tropics, the largest errors are found near the surface and in middle troposphere. The local maximum near the surface is associated to the largest background errors found near the ITCZ, for all ensembles (see top panel of figure 3.9, section 3.4.1).

The results also show that the model error simulation leads to an increase of the magnitude of the background errors in all regions, mainly in lower troposphere. However, its impact is smallest in TROPIQ region and largest in the North Pole, except in stratosphere and in upper troposphere.

In the tropics, this result may reflect that some sources of error have been underestimated. For instance, it could be important to use different deep convection parametrizations in the different ensemble members, in order to simulate the model error due to uncertainties in the treatment of deep convection, which is particularly important in this region. On the other hand, the fact that, in middle and lower troposphere, the atmospheric variability of the temperature in the tropics is relatively small (e.g., when compared with middle and high latitudes) might also partly explain the smallest contribution of the simulated model error to the background error in this region at these levels.

In the North Pole region (bottom panel), in lower troposphere, the large impact of the simulation of the model error can be related to the large uncertainties in parametrizations of turbulent heat fluxes, specially during Winter, when the PBL is frequently stably stratified in this region.

In addition, it was found that the impact of the model error is larger over land than over sea (not shown). For instance, over EUROPE (middle left panel) and ASIE (not shown) regions, when the model error is neglected the standard deviations of temperature background error are nearly constant in the vertical. However, when the model uncertainties are considered, a considerable increase of the background error is visible in the lower troposphere, largely due to resolution. The large importance of the model resolution is consistent with the presence of mountainous areas (e.g., the Alps and the Himalayas) in these regions.

Moreover, in NORAMER region, the local maximum visible in the three estimates near the PBL top is enhanced when the model error is simulated, mainly when the errors related to physical parametrization are not neglected. Possibly, these local maxima reflect a large difficulty in the prediction of the shallow clouds in this region.

It is also interesting to compare the vertical profiles shown in figure 4.12 with the profiles of RMSE of the innovation vectors for temperature (see appendix D). Nevertheless, it is important to note that an innovation inevitably contains a contribution from the observation error. This fact partly explains that the magnitude of the RMSE of the innovations is larger than the standard deviation of the background errors estimated by the ensemble approach.

The comparison between figure 4.12 (top right panel) and the left panel of figure D.1 shows that, both the ensemble derived statistics and the innovations statistics computed using radiosondes indicate that the magnitude of the background errors has maximum values in stratosphere, near the tropopause and near the surface. However, the standard deviation derived by the ensemble method presents a local maximum in middle troposphere (near 600hPa), which is not visible in the RMSE of the innovations. On the other hand, the maximum value of the RMSE of the innovations computed using aircraft reports is located near 700hPa (right panel of figure D.1).

In SUD20 region, the RMSE of innovations is maximum near the surface and decreases with

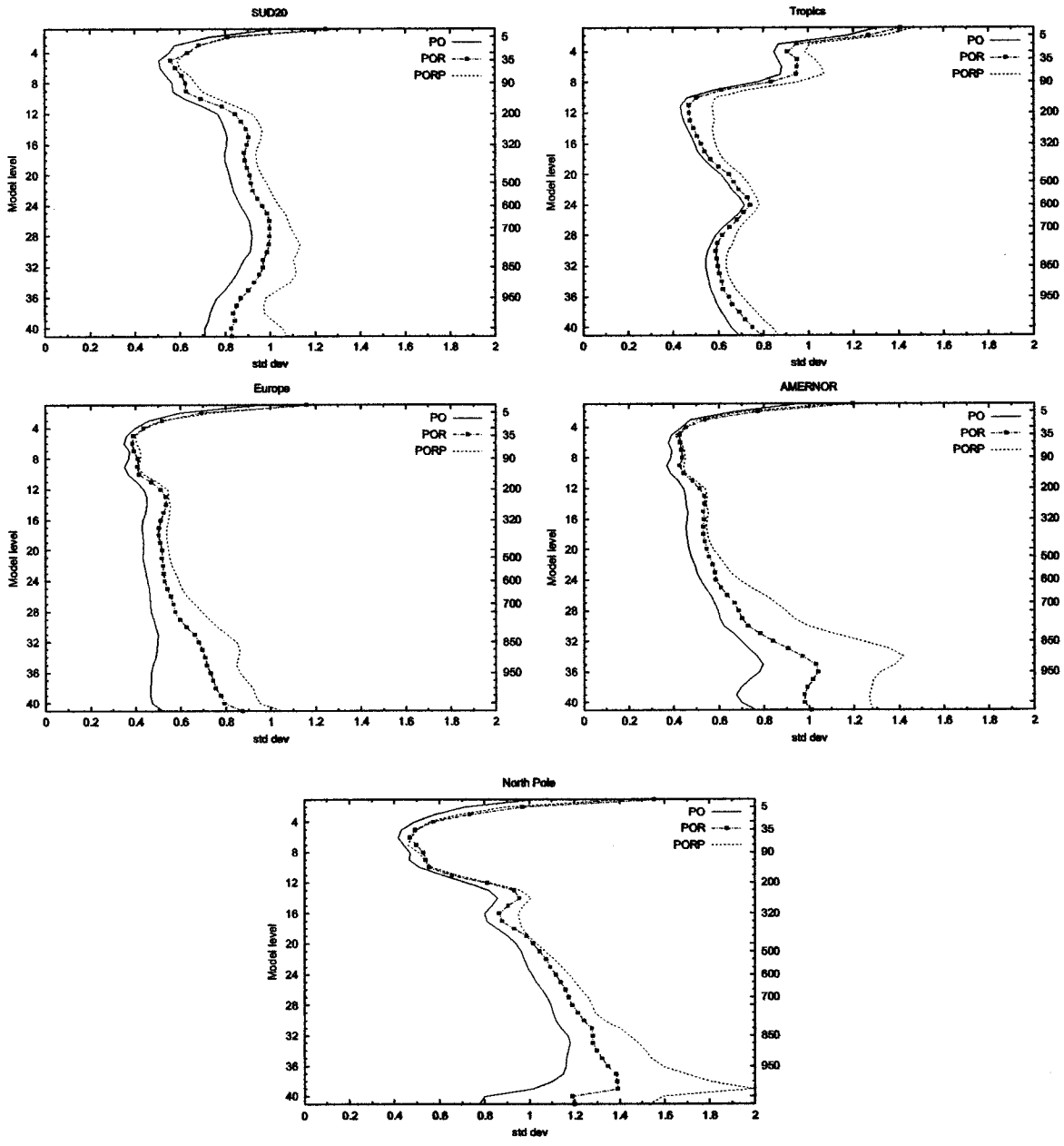


Figure 4.12: Vertical profile of standard deviation of temperature ($K$) background error estimated from PO, POR and PORP ensembles.

height, reaching its minimum value near 400hPa. On the other hand, according to the ensemble estimates, the standard deviation of background error is largest near 700hPa and it is small near the surface, mainly for PO and POR ensembles (figure D.2 and top left panel of figure 4.12).

Possibly, the low density of the observation network over the tropics and in the Southern Hemisphere partly explains the differences between the profiles of ensemble estimates and those of innovations.

Over Europe and NORD20 (not shown) regions, in middle and lower troposphere, the shape of the vertical profile of innovation RMSE is very similar to the one of standard deviation derived from PORP and POR ensembles (to contrast with PO) (figures 4.12 and D.3).

In addition, both statistics from innovations and from ensemble simulations agree that the background errors are larger in NORAMER than in EUROPE region. Moreover, the background error in NORAMER region is larger in lower troposphere than in middle troposphere, for both estimates. However, for the innovation estimates the maximum value is located at surface (figure D.4), while according to the ensemble estimates the maximum is near the PBL top.

In extra-tropical regions (as in NORAMER, EUROPE, NORD20 and SUD20 areas), the RMSE of innovations shows a maximum near the tropopause. For the ensemble derived statistics, only a small local maximum is visible in SUD20 and NORD20 regions.

These differences between the innovations and ensemble statistics near the tropopause levels might be caused by large observation errors at these levels. Another possible explanation is that some important sources of errors for the background error have been misrepresented. Nevertheless, regarding that the tropopause is higher in the tropics than in the extra-tropics and that the vertical resolution of ARPEGE decreases with height, one could expect that the uncertainties relatively to the position of tropopause would be largest in the TROPIQ region. In other words, one could expect that the background errors near the tropopause would be largest in the tropics.

The normalized standard deviations of the temperature background error were also examined for the three ensembles. The corresponding maps for level 32 (near 850hPa) are shown in figure 4.13 for PO and PORP ensembles. Some similarities between the ensemble simulations were identified. For instance, very large errors are located in the North Pacific in adjacent areas to the North American west coast. Relatively large values are also found near the West coast of the Iberian Peninsula and near the NW and SW coast of Africa. As said in the previous chapter, these results are related to the presence of the marine stratocumulus, which have a quasi-permanent character in these regions.

On the other hand, for PORP, the values over Himalayas are increased. This is mostly due to the fact that in mountainous regions, the sources of error due to resolution become more important to the background error.

In addition, it is visible that for PORP the background errors are enhanced in USA (mainly due to the uncertainties in physical parametrizations) and the local minimum in Europe is less pronounced.

Nevertheless, in the middle and high troposphere the contrasts between data rich and data poor areas are well marked for PORP, as for the other ensemble sets. This is illustrated by the normalized standard deviation of temperature background error at level 21 (near 500hPa), presented in figure 4.14 for PORP.

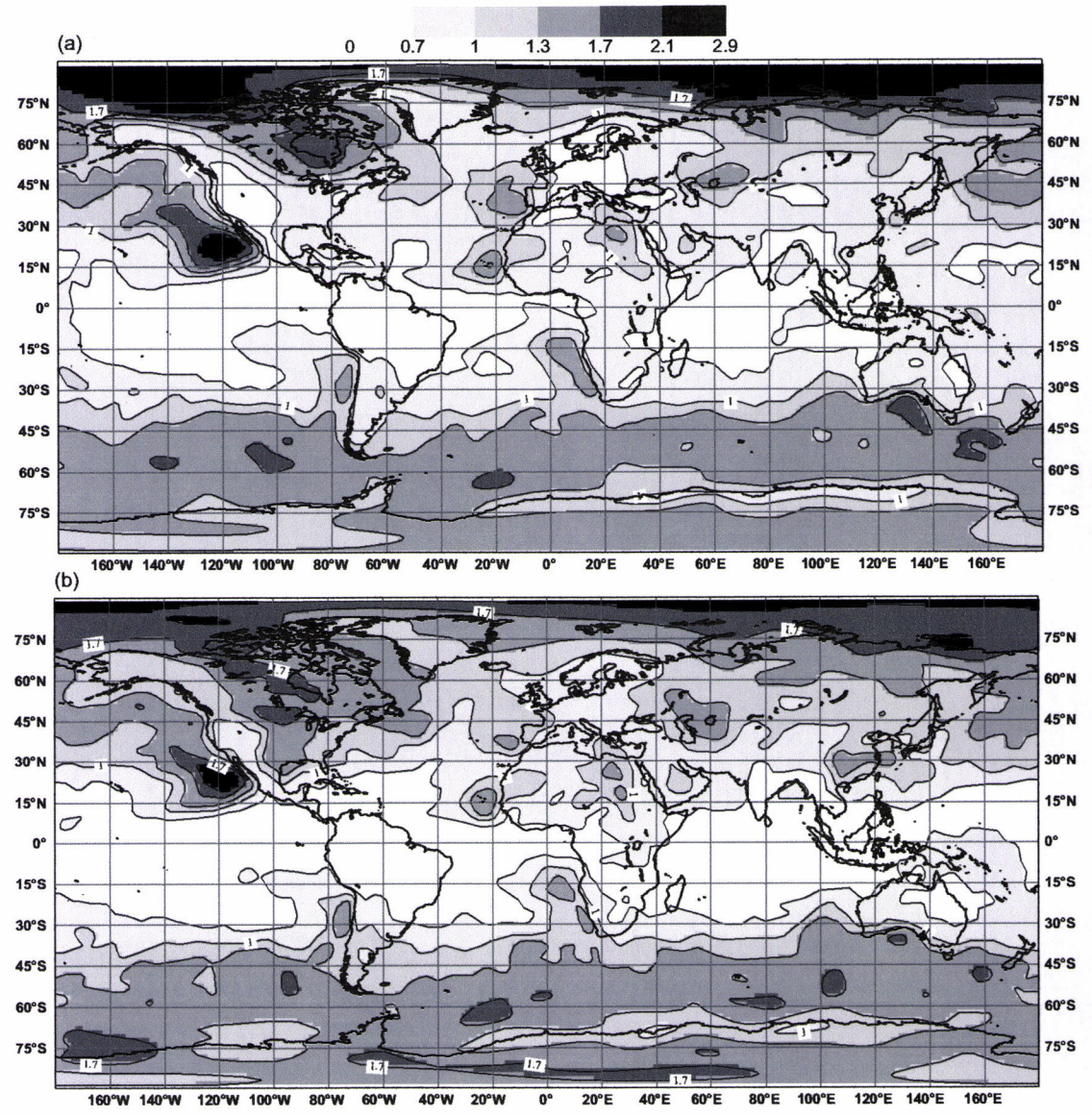


Figure 4.13: Normalized standard deviations of the temperature background error at model level 32 (around 850 hPa), for (a) PO and (b) PORP ensembles.

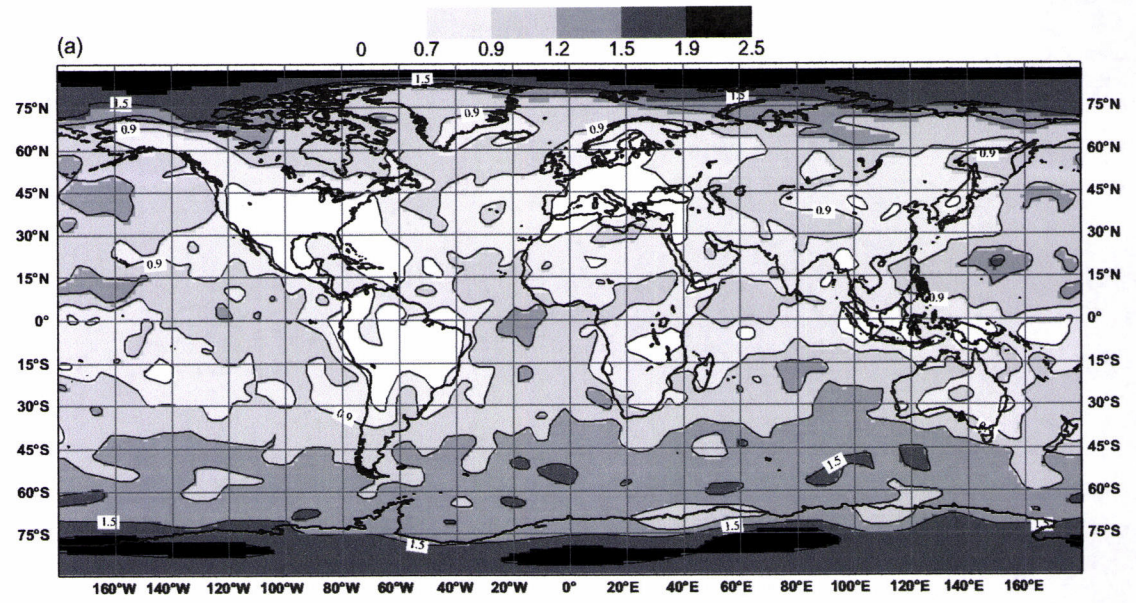


Figure 4.14: Normalized standard deviations of the temperature background error at model level 21 (around 500 hPa) for the PORP ensemble.

### 4.3.2 Horizontal length scales

The length scales for different regions (defined in table 3.1, section 3.4.2) of the globe were also examined. The corresponding vertical profiles of length scale are shown for temperature in figure 4.15. According to all ensemble simulations (see also the top panel of figure 3.14, section 3.4.2), the length scale is largest in the tropics and smallest in data rich regions (for instance, in EUROPE and NORAMER areas). However, for PORP, near the PBL top, the length scale in NORAMER region is considerably larger than for the other ensembles.

## 4.4 Sensitivity studies

In this section, the differences between the background error statistics derived from similar ensemble simulations, but at different resolutions will be presented. The experiments compared in order to accomplish this goal are presented in table 4.2.

The figures 4.16 and 4.17 present the vertical profile of standard deviation of the background error, estimated for the three types of ensembles described in table 4.2. It is visible that the vertical distribution of standard deviation is similar for both PO and POH ensemble sets. However, the magnitude of the background error is increased when estimated by an ensemble at higher resolution (POH). This is more marked for vorticity and divergence, mostly due to mesoscale phenomena, in accordance with the fact that for these variables the largest contributions to the background errors come from the mesoscales. For specific humidity, the standard deviation has also larger values in the ensemble at higher resolution (T449), mainly in lower troposphere. Moreover, the contrast between PO and POH is largest over the Northern Hemisphere (not shown). This is consistent with the larger percentage of mountainous areas in this hemisphere.

The increase of the background error magnitude when the ensemble simulations are performed with a higher resolution might look like a paradox at first sight. However, it is important to remember that the variability of the forecast fields, from a NWP model, is limited by the effective model resolution, which is defined by spectral truncation (or grid spacing) and by the

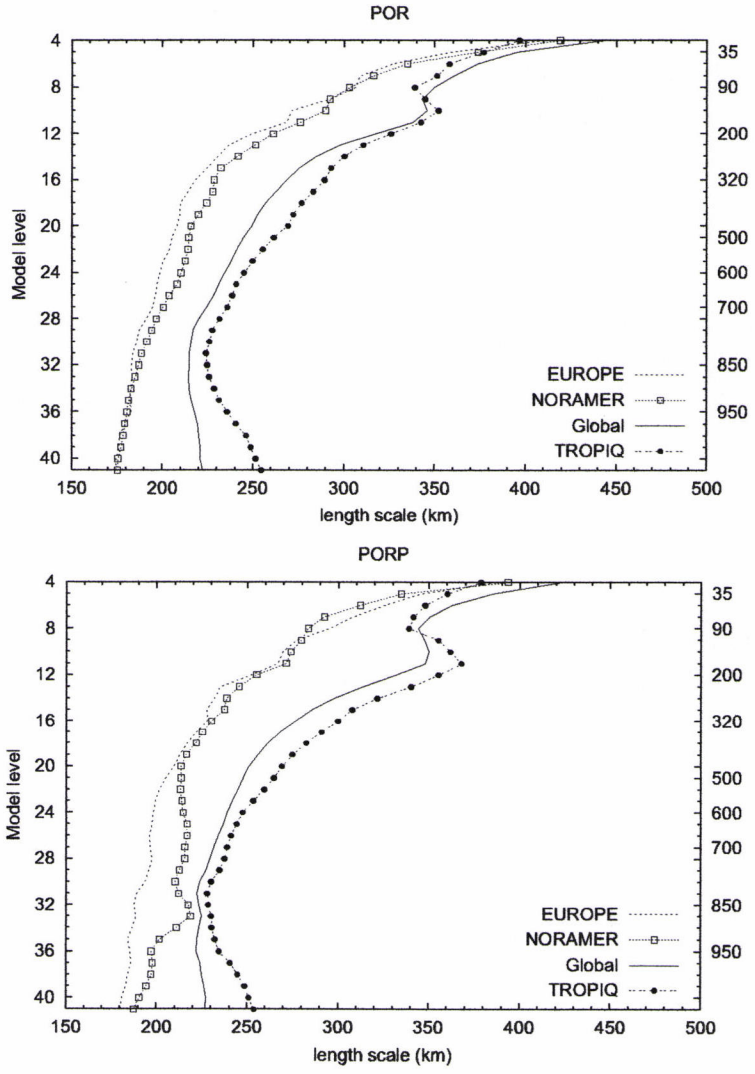


Figure 4.15: Length scales of the background error of temperature, estimated by the Ensemble Method for POR (top) and for PORP (bottom) ensembles.

Table 4.2: Ensemble simulations performed at truncation T299 and T449.

| EXP | Members | Observations perturbations | Resolution | Changes in parametrizations | Error source |
|---|---|---|---|---|---|
| PO | P33 | yes | T299 | HD(T299) | observations |
| | P34 | yes | T299 | HD(T299) | at T299 |
| POH | A16 | yes | T449 | HD(T449) | observations |
| | A12 | yes | T449 | HD(T449) | at T499 |
| POPH | A15 | yes | T449 | HD(T449)<br>vertical diffusion<br>cloudiness scheme<br>LSCON | observations<br>at T449<br>and<br>physical |
| | A12 | yes | T499 | HD(T449) | parametrizations |

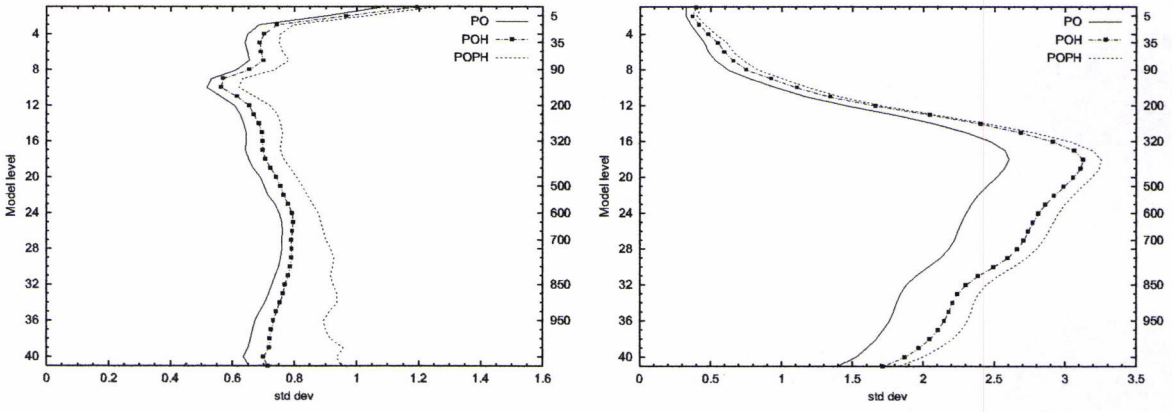


Figure 4.16: Vertical profile of standard deviation of temperature ($K$) (left panel) and vorticity ($10^5 s^{-1}$) (right panel) background errors estimated by Ensemble methods.

smoothing due to parameterized horizontal diffusion. In other words, as the effective resolution of the model decreases, its variability at smaller scales will be smaller, because the smaller scales are either not resolved at all, or dissipated by the horizontal diffusion.

In order to evaluate the implications of these results, it is useful to remember that the scales that are resolved by the observation network, but not by the model due to its limited resolution, are treated as a part of the observation representativeness errors (rather than a part of the background errors). This is the case in particular for scales which are not represented at all by the model. More generally, one may wonder if, the small scale variability that is dissipated by the numerical diffusion should also be considered and treated as a part of these representativeness errors.

For the above reasons, the determination of an optimum amount of small scale variance for data assimilation remains an open issue. On one hand, using relatively large small scale variances may be seen as realistic and useful, as an attempt to correct (through data assimilation) what may be considered as a model deficiency. On the other hand, this might also be seen as allowing observations to introduce many small scale features, which may be considered as noise by the model during the forecast integration, and therefore should rather be filtered by the analysis.

A possible strategy may then be at least to make the model diffusion as realistic as possible,

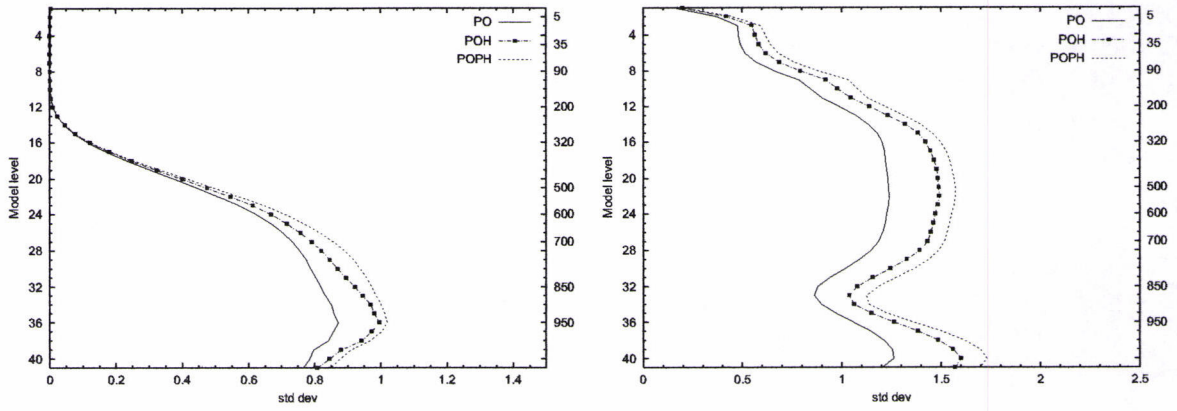


Figure 4.17: Vertical profile of standard deviation of specific humidity ($g\,kg^{-1}$) (left panel) and divergence ($10^5 s^{-1}$) (right panel) background errors estimated by Ensemble methods.

and then to derive some background error statistics that are consistent with this model diffusion.

From the comparison between POH and POPH, one can see the role of model error due to parametrizations uncertainties at a higher resolution (T449). It is visible that for POPH the magnitude of vorticity and divergence background errors is increased uniformly in all troposphere. However, for specific humidity, the standard deviation for POPH are only sightly larger than those for POH.

For temperature, the enhancement of the background errors due to perturbations of the physical parametrizations is more pronounced in the lower and middle troposphere. Moreover, the examination of variance and correlation spectra shows that this increase of the background error magnitude is mainly due to the synoptic scale phenomena and that the contributions from mesoscales are less emphasized (not shown). As a consequence, the temperature correlation length scales are larger for POPH than for the other two ensembles sets. These results are in agreement with those found for the ensembles at T299.

The figure 4.18 shows the correlation spectra of vorticity and specific humidity background error at level 32 (near 850hPa). It is visible that the contributions from the small scales are emphasized for both POH and POPH ensemble simulations. As a consequence, the correlation length scales are smaller than for PO (not shown). The same result was found for divergence. In addition, in lower troposphere, the contributions of the unbalanced components in the multivariate formulation are emphasized, mainly for the POPH ensemble. This is illustrated in figure 4.19 for temperature.

Moreover, the comparison between figures 4.6 and 4.19 shows that the contributions of the balanced components are more pronounced for POPH than for PORP. In other words, the explained variance by the balance relationships is larger when the ensemble is built from models having the same resolutions (POPH) than when models with different resolution have been used in the ensemble simulations (PORP).

The figure 4.20 presents the standard deviation of the background error of $P_s$ as a function of the wave number (or horizontal scales). It may also be mentioned that for the POH experiment there is less diffusion of the small scale features than in PO experiment. The comparison between PO and POH suggests that the large scale phenomena are better represented at higher resolution (POH), while the uncertainties in mesoscales are enhanced. This suggests that if the smaller scales are better represented by the model this has a positive impact on the large scales. This is consistent with classical concept of atmospheric predictability, according to which the errors in smaller scales tend to propagate to the larger scales.

Moreover, it is visible that the simulation of model errors due to uncertainties in the

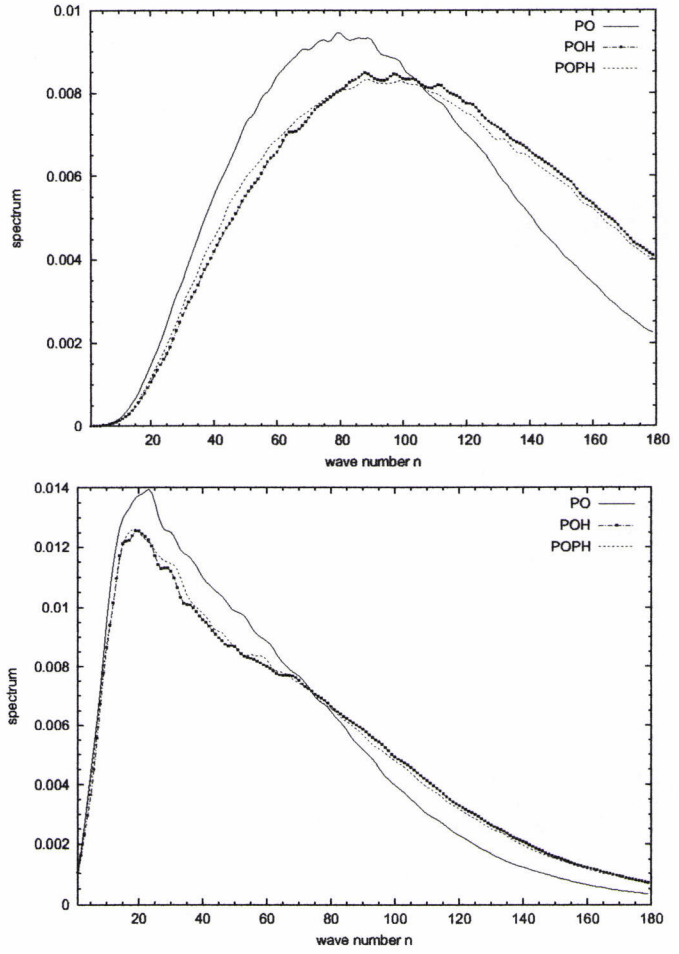


Figure 4.18: Correlation spectra of vorticity (top) and specific humidity (bottom) background error at level 32 (near 850hPa).

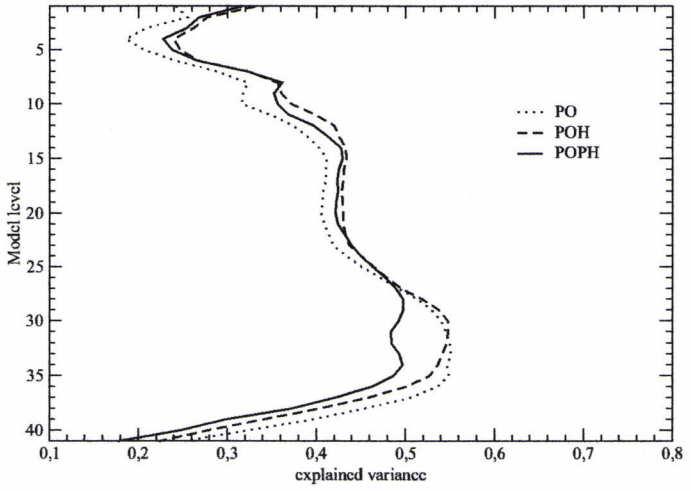


Figure 4.19: Explained variance for temperature background error, using the PO, POH and POPH ensembles.

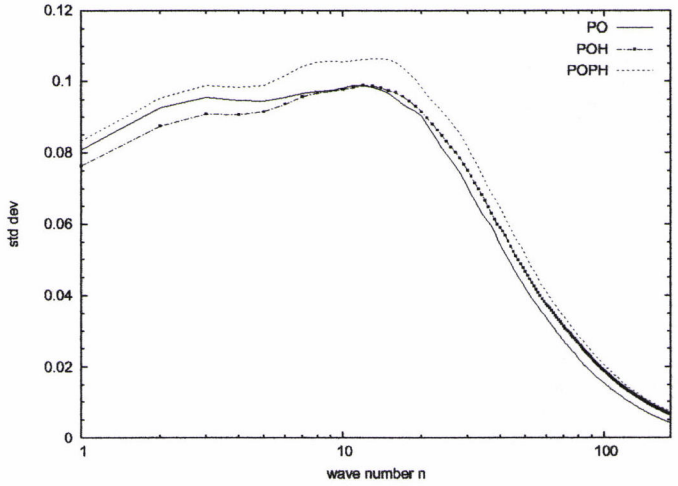


Figure 4.20: Standard deviation of surface pressure background error as function of the wave number, estimated by the Ensemble method, using the PO, POH and POPH ensembles.

parametrizations contributes to increase the magnitude of surface pressure background error, mostly in the synoptic scales (for wave numbers between 4 and 20).

## 4.5 Impact on the forecast scores

Some assimilation and forecast experiments have been performed in order to investigate the impact on the forecast quality of the statistics provided by the ensemble approach when the model error is simulated. The forecast scores were thus compared with those from the ensemble method in the perfect-model context (which were operational when these impact experiments were done). For this purpose, both the impact of the global covariances and of the local standard deviations have been studied.

The experiments were done in a similar way to those presented in the previous chapter. The list of experiments and their features are summarized in table 4.3. These experiments were performed for the period between the 1st and the 28th of February 2005.

Table 4.3: Impact Experiments. The following letters are used to refer to the different statistics: OPE = ensemble statistics using the perfect-model assumption, LSD = ensemble local standard deviations provided by the PORP ensemble experiments, TPH and TPR correspond to the ensemble global covariances provided respectively by the POPH and PORP ensemble experiments. The 2nd column refers to the global covariances. The 3rd column refers to the local standard deviations that are used in the minimization (min.). The last column refers to the value of $R_{NMC}$ used.

| Experiment | Global covariances | Local std. dev. (min.) | $R_{NMC}$ |
|---|---|---|---|
| OPE | PO | PO | 1.5 |
| LSD | PO | PORP | 1.5 |
| TPH | POPH | PO | 1.3 |
| TPR | PORP | PO | 1.3 |

The RMSE and bias scores were calculated by comparing the forecasts with both ECMWF

analysis and TEMP observations, for geopotential, wind, temperature and relative humidity. The impact of simulating the model error in the ensemble approach was mostly neutral for humidity and wind. The largest impact on the scores was found for geopotential, mainly when compared with TEMP observations. Therefore, the geopotential scores with respect to the TEMP observations will be shown.

As mentioned in section 4.2.1, the simulation of the model error in the ensemble approach leads to an increase of the global standard deviations. Therefore, the value of $R_{NMC}$ is smaller for the TPH and TPR experiments than for the OPE experiment, which provides similar temperature standard deviations for the three experiments (they are still slightly larger for TPH than for OPE).

The left panels of figures 4.21 and 4.22 correspond to the impact of using the ensemble global covariances provided by the POPH experiment (where the model error due to uncertainties in the parametrizations is simulated at high resolution), when compared with those provided by the PO experiment. The impact is negative in the stratosphere and near the tropopause, over SUD20, TROPIQ and AUS/NZ domains. This is illustrated for the bias on the last two domains in figure 4.21, and for the RMSE on TROPIQ in figure 4.22.

In the troposphere, the scores are generally neutral except over the AUS/NZ region (not shown), where the RMSE scores are positive (also when compared with the ECMWF analysis).

The right panels of figures 4.21 and 4.22 correspond to the impact of using the ensemble global covariances provided by the PORP experiment (where the model error associated to resolution and to uncertainties in the parametrizations is simulated), when compared with those provided by the PO experiment. The impact seems to be positive over the tropics (essentially for bias, while the impact is neutral on standard-deviation (not shown)). Moreover, a reduction of the geopotential bias is visible over the AMNORD and NORD20 domains in the stratosphere and over AUS/NZ domain in the upper and middle troposphere. In the other regions, the stratospheric impact is globally neutral. These positive impacts might be caused by the relaxation of the mass/wind coupling in the PORP experiment (when compared with PO and POPH experiments, see figures 4.6 and 4.19).

However, in the middle and upper troposphere, the RMSE scores are neutral or negative. The strongest negative impact is found in the AMNORD region (right panel of figure 4.22), which is also visible on wind (not shown). This is likely to be due to the reduction of geostrophic coupling verified in the PORP experiment.

For temperature, the scores are mostly neutral, except in PBL levels, where there is a positive impact, mainly in AMNORD and AUS/NZ regions. This may be explained by the increase of the background error variance (in particular at these levels) due to the model error simulation.

It was also found that the impact of the new $\sigma_b$ map (LSD experiment) in the minimization is mostly neutral. This is consistent with the relative closeness between the two maps of ensemble standard deviations, in particular for vorticity. It may be mentioned that the two maps of ensemble standard deviations are more different for temperature, but that they are not used directly in the Arpége 4D-Var (the vorticity maps are used instead, to determine the main part of the wind and temperature covariances).

## 4.6 Summary

In the present chapter, the estimates of the background error covariances from the ensemble method when the model error is simulated are examined. The model error is simulated by performing an ensemble of experiments using different resolutions and different values for certain tuning parameters from some physical parametrizations. The resulting covariances are compared with those provided by the ensemble approach in a perfect-model context. Furthermore, the

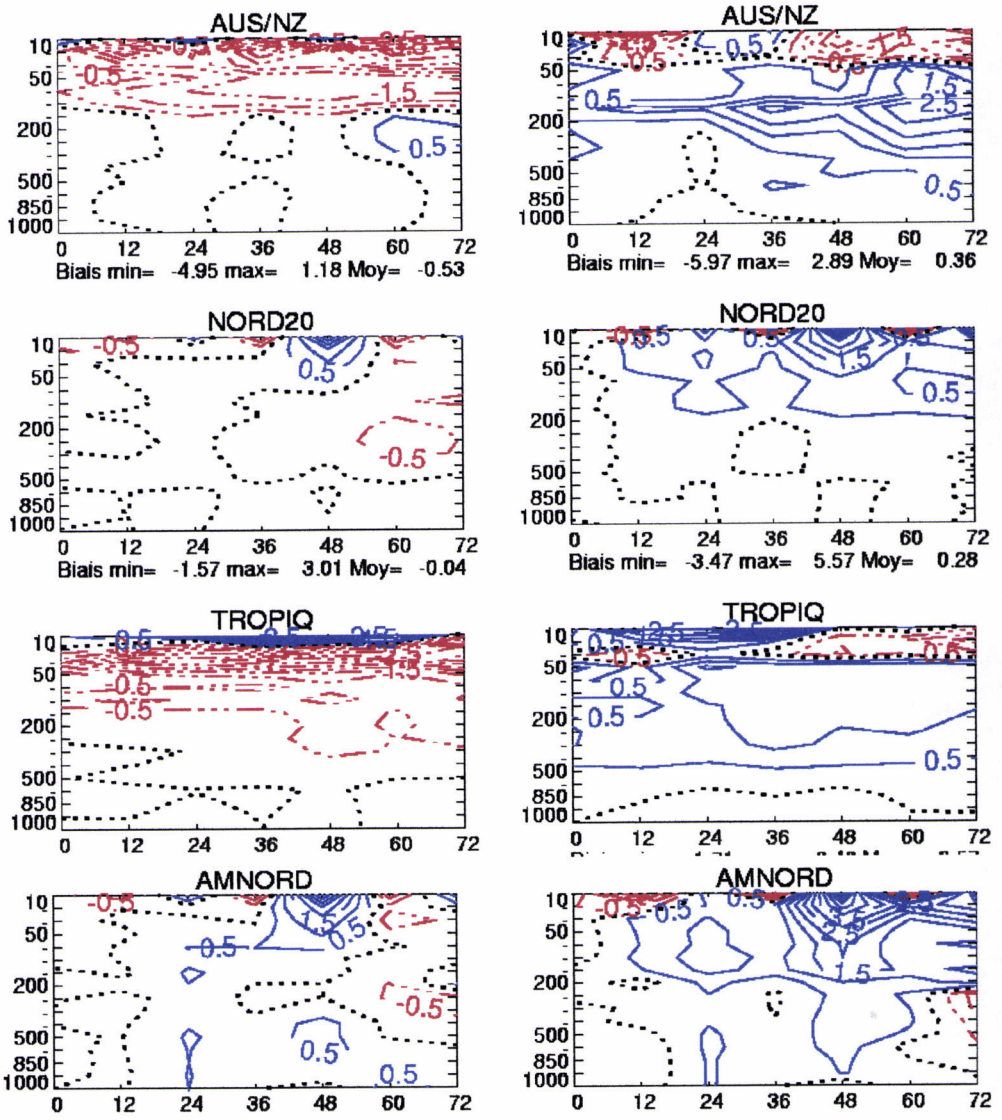


Figure 4.21: Difference between the bias (ME, in m) of geopotential forecasts (as function of pressure levels and forecast range) of the OPE experiment and those of the following experiments: TPH (left panels) and TPR (right panels). Forecasts are verified against the TEMP observations. Blue solid (resp. dotted) isolines indicate ranges and levels, for which the scores of the involved experiment are better than (resp. similar to) the scores of the OPE experiment. Red dashed-dotted isolines indicate ranges and levels, for which the scores of the involved experiment are worse than the scores of the OPE experiment. The isoline spacing is 0.50 m. The acronyms of the 3 areas are defined in table 3.1.

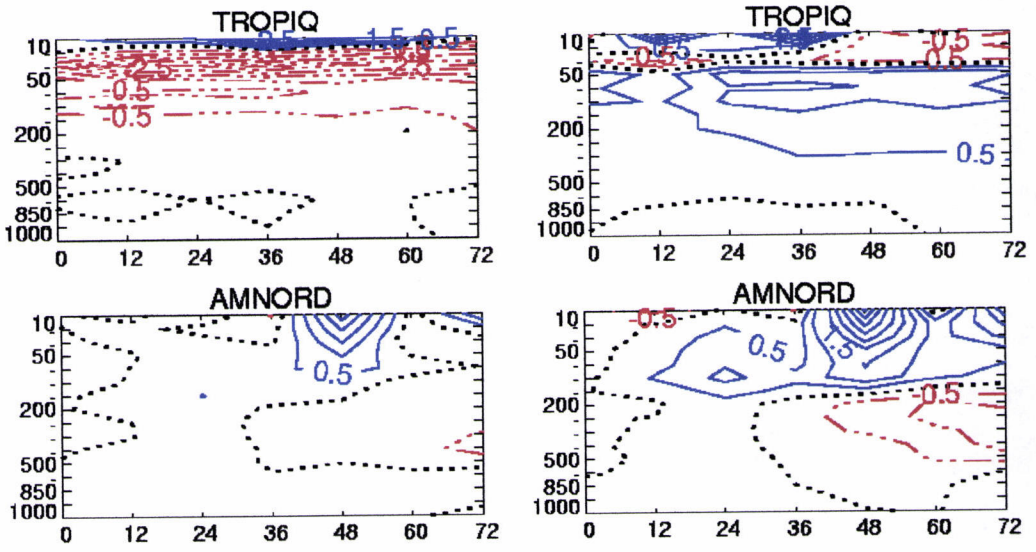


Figure 4.22: The same as in figure 4.21, but for the RMSE (m) in the TROPIQ and AMNORD regions.

impact of the corresponding covariance estimates on the forecast quality was also studied.

The effect of the model error simulation in the ensemble approach is to increase the magnitude of the estimated background error variance. The results indicate that the resolution is the most important contribution for small scale variables (e.g., vorticity and divergence). The perturbations of the physical parametrizations remain important in the tropics for vorticity and divergence, and also for temperature in general. This is more obvious in lower and middle troposphere in the extra-tropics and in upper levels over the tropics.

Furthermore, for the large scale variables (temperature and surface pressure) the increase in the magnitude of the background errors is associated to the enhancement of the synoptic scale contributions. On the other hand, for the other variables, the simulation of the model error in the ensemble approach appears to emphasize the relative contributions of the mesoscales to the background errors. As a consequence, there is a reduction of the coupling between mass and wind variables, mainly due to the simulation of errors related to resolution.

This reduction of the mass/wind coupling has a positive impact on the forecast quality of geopotential over the tropics and in some stratospheric extra-tropical areas. On the other hand, the reduction of mass/wind balance leads to a degradation of the forecasts of geopotential and wind in some tropospheric extra-tropical regions. This suggests that the mass/wind coupling is smaller in the tropics and in the stratosphere, and that the relaxation of this coupling as a function of latitude and height should be strengthened in the $\mathbf{B}$ matrix.

It was also found that there is a better agreement between the vertical profiles of the statistics of the innovations and those from the ensemble approach when the model error is simulated.

The results also show that the magnitude of the background error increases when the horizontal resolution of model used in the ensemble simulations increases. This is more pronounced for small scale variables, for which the largest contributions to the background errors are due to the mesoscale phenomena.

The increase of the background error magnitude when the ensemble simulations are performed with a higher resolution may look like a paradox at first sight. However, it is important to remember that as the effective model resolution decreases, its variability at smaller scales will be smaller, because the smaller scales are either not resolved at all or dissipated by the horizontal diffusion.

Moreover, small scale structures that are observed but not represented at all by the model are usually considered and treated as a part of the observation representativeness errors, rather than a part of the background errors. Thus, one may wonder if the small scale variability that is dissipated by horizontal diffusion should also be considered and treated as a part of these representativeness errors. This suggests that the determination of the optimum amount of small scale variance remains then an open issue, for the specification of the **B** matrix in data assimilation.

Finally, it was found that, for surface pressure, the large scales are better represented when using a higher resolution. This suggests that an improved representation of smaller scales by the model has a positive impact on the large scales. This is consistent with a classical concept of atmospheric predictability, according to which the errors in smaller scales tend to propagate to the larger scales.

# Chapter 5

# The impact of neglecting the observation errors in the ensemble approach

## 5.1 Introduction

The results presented in the previous chapters have shown that there is valuable information in the background error covariances (locally and globally) estimated by the ensemble approach in a perfect-model context and when the model error is simulated.

In the present chapter, the effect of neglecting the observation error in the ensemble method is investigated. In this case, it is assumed that the background error is caused mainly by the model error. Formally, this corresponds to setting $\delta_o(n) = 0$, for any ensemble member $n$ (see section 2.3). The resulting covariances are compared with those provided by the ensemble approach in a perfect-model context, and with those from the ensemble method when both the observation and the model errors are simulated. The results presented here are focussed on the spectral space (at truncation T179).

## 5.2 Neglecting observation errors when perturbing the physics

In order to study the effect of neglecting the observation errors in the ensemble method, firstly it is supposed that the background error is mostly due to model errors associated to the physical parametrizations. The list of experiments analysed to accomplish this aim are described in table 5.1. This section summarizes the main results obtained from the comparison between the background error covariances provided by the three types of ensembles presented in this table.

The figures 5.1 and 5.2 present the vertical profile of standard deviations of background errors, estimated by ensemble method for the three types of ensemble differences described in table 5.1. It is clear that the magnitude of the background errors is largest for POPH, as expected, since both the observation errors and model error due to physical parametrizations are simulated.

In addition, one can notice that for temperature in troposphere there are significant differences between the ensemble sets. For instance, according to POH, the magnitude of background errors of temperature has small variations in vertical and has its maximum in middle troposphere. On the other hand, for PPH and POPH, the magnitude of the background errors is maximum near the surface and decreases with height (mainly for PPH).

Table 5.1: Ensemble experiments performed to study the impact of neglecting the observation error. Here, HD is the horizontal diffusion and LSCON is the large scale convergence used in the parametrization of convection. These experiments were performed at truncation T449.

| EXP | Member | Observations perturbations | Changes in parametrizations | Error source |
|---|---|---|---|---|
| POH | A16 | yes | HD(T449) | observations |
| | A12 | yes | HD(T449) | |
| PPH | A14 | no | HD(T449)<br>vertical diffusion<br>cloudiness scheme<br>LSCON | physical<br>parametrizations |
| | A11 | no | HD(T449) | |
| POPH | A15 | yes | HD(T449)<br>vertical diffusion<br>cloudiness scheme<br>LSCON | observations<br>and<br>physical |
| | A12 | yes | HD(T449) | parametrizations |

For vorticity, all ensembles agree that the magnitude of vorticity background errors increases significantly with height, reaching its maximum near the tropopause. This maximum is related to the jet streams, as said before. Nevertheless, the contrast between the tropopause levels and the low levels is smallest for PPH (when the observations errors are not simulated).

Moreover, for specific humidity and divergence, as for vorticity, the magnitude of the background error is smallest for PPH and largest for POPH, at all levels. This result might suggest that the uncertainties related to the observations have a larger contribution to the background error than the model error. On the other hand, this result may also suggest that some important sources of model error have been underestimated.

The comparison of the correlation spectra from the POPH and POH estimates, for $\zeta$, $\eta$ and $q$, suggests that the effect of neglecting the model error (due to uncertainties in physical parametrizations) is neutral. In addition, the comparison between the POPH and PPH estimates indicates that when the observation errors are neglected there is a decrease of the relative contribution of the background error associated to the large scales, and the mesoscale contributions are emphasized. This is illustrated in figure 5.3 for specific humidity at model level 32 (near 850hPa). As a consequence, for $\zeta$, $\eta$ and $q$, the corresponding correlation length scales are smaller for PPH than for POH and POPH (not shown).

The figure 5.4 presents the correlation spectra for temperature background errors (at model level 32, near 850hPa). It is visible that when either the observation errors or either model errors are neglected, the synoptic scale contributions to the temperature background errors are less emphasized. It is also obvious that when the model error is simulated and the observation errors are neglected (PPH), there is a reduction of the synoptic and planetary scales contributions to background error, and the mesoscale contributions are enhanced, when compared to the POH estimate. As shown in figure 5.5, this implies that the corresponding correlation length scales of the temperature background error are smallest for PPH and largest for POPH.

The figure 5.6 presents the standard deviation of the background error of surface pressure as a function of the wave number (or horizontal scales). Is is clear that neglecting the observation

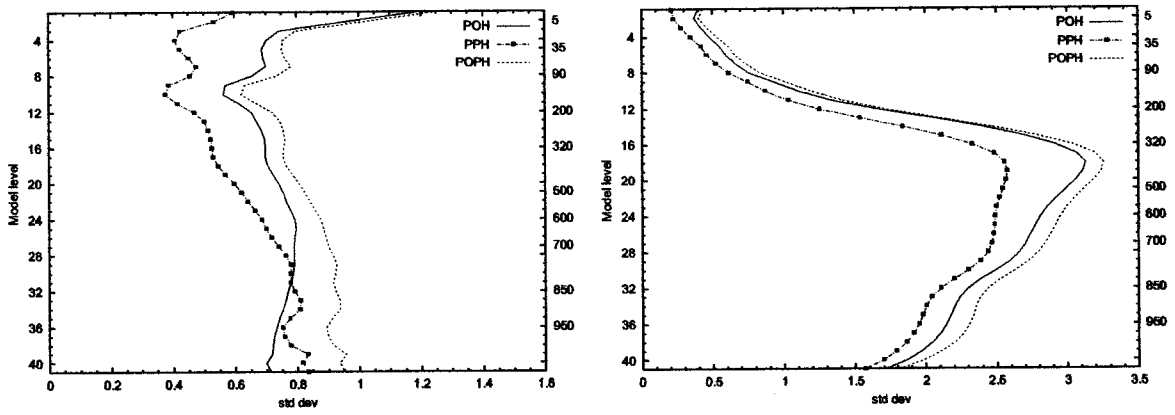


Figure 5.1: Vertical profile of standard deviation of temperature (K) (left side) and vorticity ($10^{-5}s^{-1}$) (right side) background errors, estimated by the ensemble method for the experiments described in table 5.1.

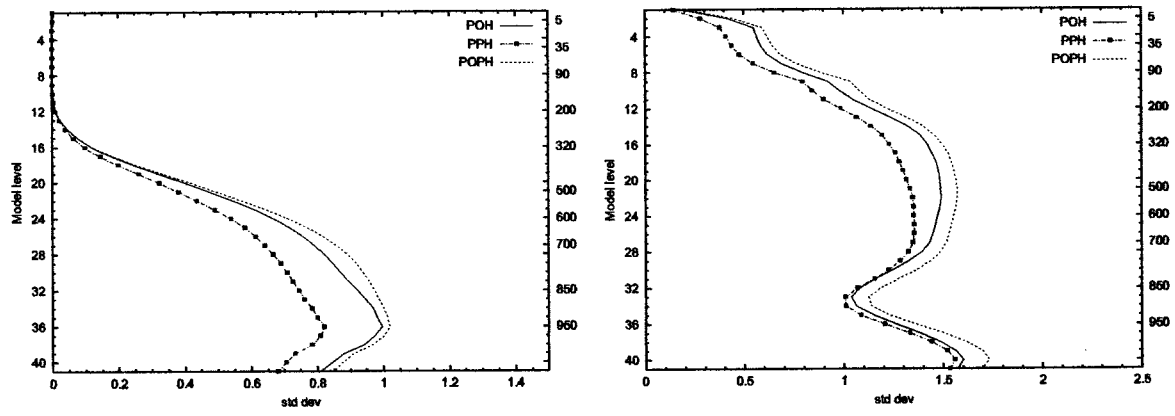


Figure 5.2: Vertical profile of standard deviation of specific humidity (K) (left side) and divergence ($10^{-5}s^{-1}$) (right side) background errors, estimated by the ensemble method for the experiments described in table 5.1.

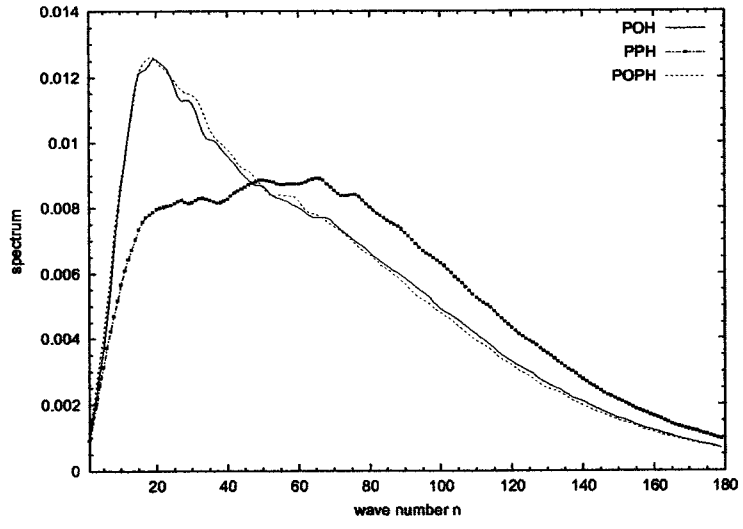


Figure 5.3: Correlation spectra of the specific humidity background error at model level 32 (near 850hPa), estimated using the POH, PPH and POPH ensembles.

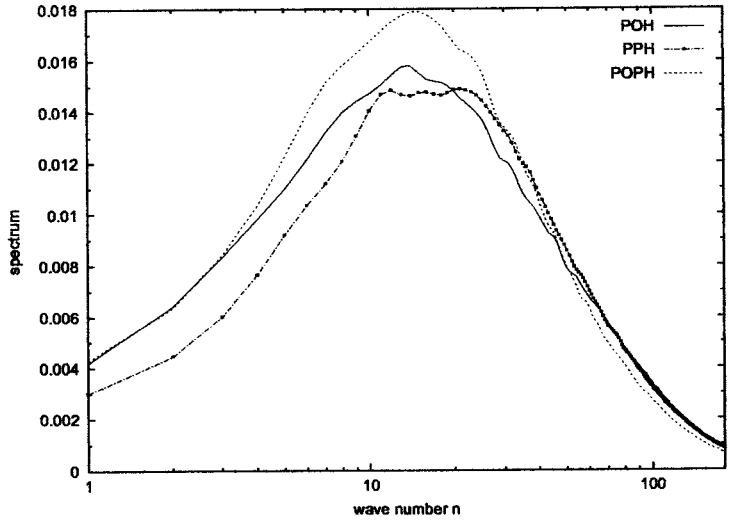


Figure 5.4: Correlation spectra of the temperature background error at model level 32 (near 850hPa), estimated using the POH, PPH and POPH ensembles.

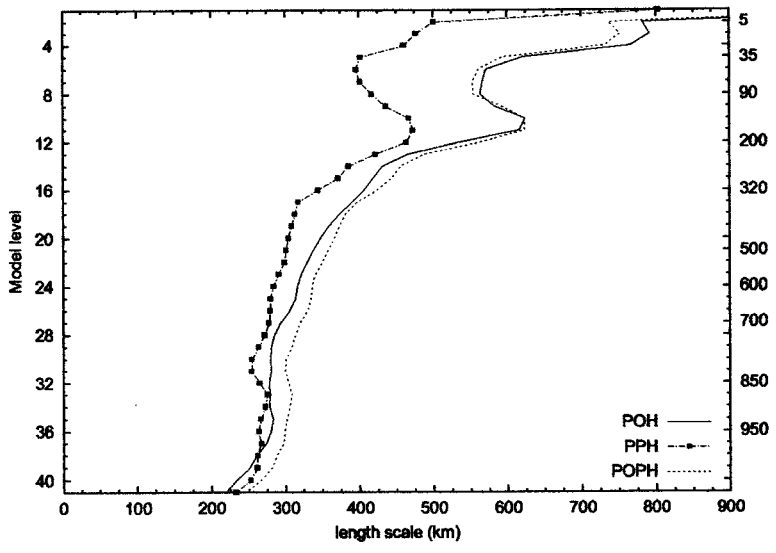


Figure 5.5: Horizontal length scale of the autocorrelation function of the temperature background error, estimated using the POH, PPH and POPH ensembles.

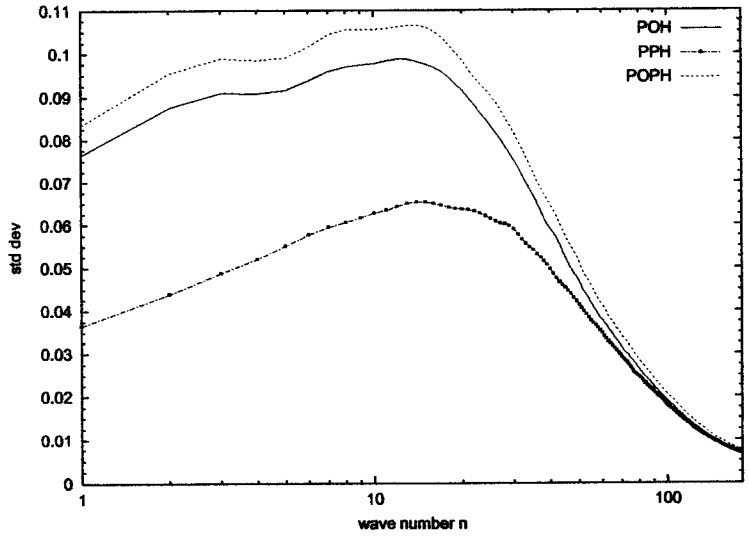


Figure 5.6: Standard deviation of surface pressure background error, estimated using the POH, PPH and POPH ensembles.

errors or the model error leads to a decrease of the variance of background errors, associated to planetary and synoptic scales (when compared with POPH). However, this decrease is considerably stronger when the observation errors are neglected (in PPH) than when the model error is not simulated (in POH). This result suggests that neglecting the observation errors in the ensemble approach might lead to an underestimation of the large scale contributions to the background errors.

Moreover, the reduction of the large scale background error variance in PPH (compared to POPH or POH) is consistent with the filtering properties of the analysis. Remember that the analysis equation applied to the evolution of errors $e$ is the same as the equation of evolution of the ensemble state differences $\varepsilon$ (compare equations 2.2 and 2.4). Consequently, the simulated analysis error $\varepsilon_a$ can be seen as the combination of a large scale part of the observation differences $\varepsilon_o$ and of a small scale part of the background differences $\varepsilon_b$ (see sections 1.3.4, 2.2 and 2.3).

Therefore, when the observation differences $\varepsilon_o$ are set to zero, the analysis differences $\varepsilon_a$ will tend to correspond simply to a small scale part of the background differences $\varepsilon_b$. In other words, due to the absence of the large scale contribution $\mathbf{K}\varepsilon_o$, the simulated analysis error tends to be relatively small scale. Moreover, the same feature may be expected for $\varepsilon_b$, because they are directly related to the simulated analysis errors.

## 5.3 Neglecting observation errors when perturbing the physics and the resolution

In this section, the effect of neglecting the observation errors in the ensemble method is studied by supposing that the background error is mostly due to uncertainties related to resolution and to physical parametrizations. The resulting background error covariances are compared with those provided by two other ensemble sets (see table 5.2). This section summarizes the main results obtained from this comparison.

The figure 5.7 presents the vertical profile of standard deviations of temperature and vorticity background errors, estimated by ensemble method for the three types of ensemble differences described in table 5.2.

For temperature, in troposphere, significant differences between the ensemble sets are visible.

Table 5.2: Ensemble experiments performed to study the impact of neglecting the observation errors. Here, HD is the horizontal diffusion and LSCON is the large scale convergence used in the parametrization of convection.

| EXP | Member | Observations perturbations | Resolution | Changes in parametrizations | Error source |
|---|---|---|---|---|---|
| PO | P33 | yes | T299 | HD(T299) | observations |
| | P34 | yes | T299 | HD(T299) | |
| PRP | A14 | no | T449 | HD(T449)<br>vertical diffusion<br>LSCON<br>convection | resolution<br>and<br>physical |
| | P30 | no | T299 | HD(T299) | parametrizations |
| PORP | A15 | yes | T449 | HD(T449)<br>vertical diffusion<br>cloudiness scheme<br>LSCON | observations,<br>resolution<br>and<br>physical |
| | P35 | yes | T299 | HD(T299) | parametrizations |

For instance, according to PO, the magnitude of background errors of temperature has small variations in vertical and has its maximum in middle troposphere. On the other hand, for PRP and PORP, the magnitude of the background errors is maximum near the surface and decreases with height. Nevertheless, the contrast between the tropopause levels and the low levels is weaker for PORP than for PRP.

For vorticity, near the tropopause, the magnitude of background errors is considerably larger for PORP (when the observations and model are perturbed) than for PRP and PO experiments. It also interesting to notice that the contrast between the tropopause levels and the low levels is smallest for PRP (when observation errors are neglected).

These results might indicate that the influence of observation errors is particularly important in the upper troposphere.

For divergence, as for vorticity, the magnitude of the background error is larger for PRP than for PO, mainly in lower and middle troposphere. The comparison between these vertical profiles and those presented in the previous section indicates that the uncertainties associated to resolution are the main cause for the differences between PRP and PO. This result is consistent with the ones presented in the previous chapter.

For specific humidity, the magnitude of the background error is slightly smaller for PRP than for PO. In addition, if both the model error and observation errors are simulated (PORP), the standard deviations are enlarged (figure 5.8).

The results also show that when the model error is simulated (PRP and PORP) the large scale relative contributions to the background error of $\zeta$, $\eta$ and $q$ are reduced (and the mesoscales are enhanced). This is particularly marked when the observation errors are neglected. This is illustrated in figure 5.9 for vorticity at model level 16 (near 300hPa). As a consequence, the corresponding correlation length scales are smallest for PRP and largest for PO (not shown).

The figure 5.10 presents the standard deviation of the background error of surface pressure as a function of the wave number (or horizontal scales). It is visible that neglecting either the observation errors or the model error leads to a decrease of the variance of background

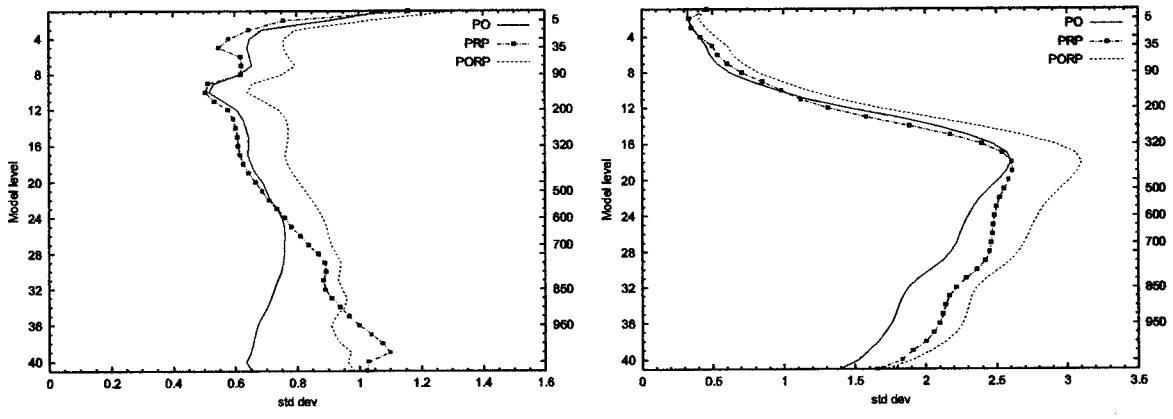


Figure 5.7: Vertical profile of standard deviation of temperature (K) (left side) and vorticity ($10^{-5}s^{-1}$) (right side) background errors, estimated by the ensemble method for the experiments described in table 5.2.

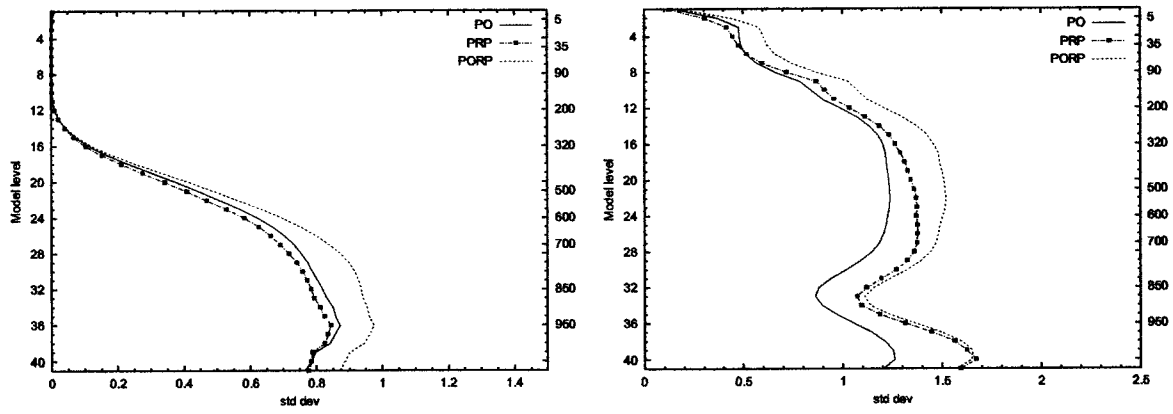


Figure 5.8: Vertical profile of standard deviation of specific humidity (K) (left side) and divergence ($10^{-5}s^{-1}$) (right side) background errors, estimated by the ensemble method for the experiments described in table 5.2.

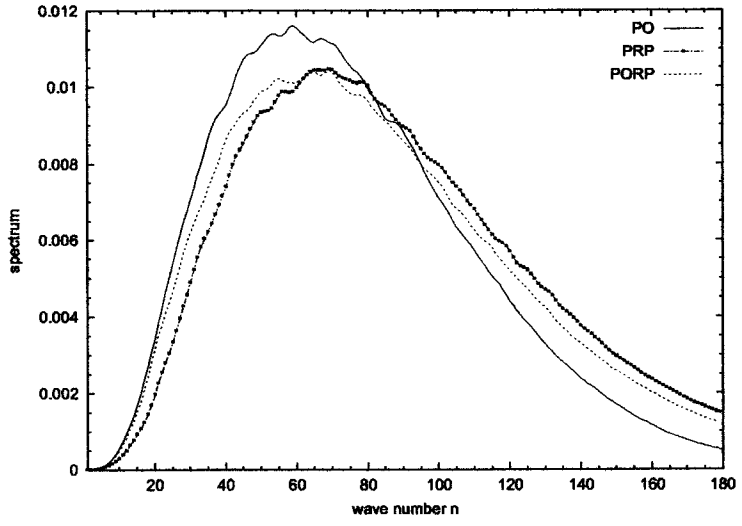


Figure 5.9: Correlation spectra of the vorticity background error at model level 16 (near 300hPa), estimated using the PO, PRP and PORP ensembles.

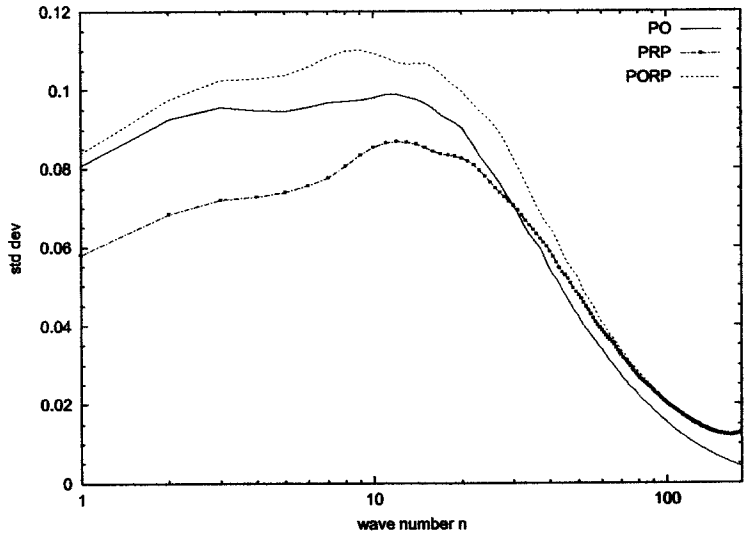


Figure 5.10: Standard deviation of surface pressure background error as a function of the wavenumber, estimated using the PO, PRP and PORP ensembles.

errors, associated to planetary and synoptic scales (when compared with PORP). However, this decrease is considerably stronger when the observation errors are neglected. As mentioned in section 5.2, this is consistent with the filtering properties of the analysis.

## 5.4 Summary

In the present chapter, the background error covariances estimated by the ensemble approach in a perfect-observation context have been described. In this framework, it is assumed that the background error is caused mainly by the model error (either due to uncertainties in physical parametrizations and resolution, or only due the uncertainties in parametrizations). In addition, the resulting covariances are compared with those provided by the ensemble approach in a perfect-model context and with those from the ensemble method when both the observation and the model errors are simulated.

Globally, it was found that neglecting the observation errors might lead to an underestimation of the background error variance. The assumption that observations are perfect leads to a decrease of the background error variance, which is particularly large in upper troposphere and for surface pressure. Moreover, this decrease is mostly due to a reduction of the large scale contributions to the background error (in accordance with the expected filtering properties of the analysis).

On the other hand, in agreement with the previous chapter, the results also suggest that neglecting the model error might also cause an underestimation of the background errors, mainly in lower troposphere. In particular, the uncertainties due to physical parametrizations appear to be not negligible for the temperature background errors, while the resolution seems to be more important for background errors of vorticity and divergence.

# Chapter 6

# Conclusions and perspectives

The estimation of the background error statistics is a key issue for data assimilation. In the last years, the global covariances used in the Arpège 4D-Var, as in other NWP models, have been determined by using the NMC method. More recently, an interesting alternative technique based on an ensemble of analysis experiments was proposed by Houtekamer *et al.* (1996) and Fisher (2003). In the present work the ensemble approach has been applied to the nonstretched version of the Arpège global model.

First, it was assumed that the forecast model used in the ensemble simulations is perfect. The corresponding estimates of the covariances were compared with those that were operational, and in particular with those provided by the NMC method.

Globally, the ensemble approach appears to emphasize the relative contributions of the small scales to the background errors. Consequently, the correlation functions are sharper in the ensemble method than in the NMC method. These results are in accordance with those described by Fisher (2003), where these differences were considered to be caused by the involvement of longer forecast ranges in the NMC technique.

Nevertheless, the analysis error estimated by the NMC method, strongly depends on the analysis increment, and it was found that its spectrum is much larger scale than the analysis dispersion spectrum from the ensemble method. Moreover, this result is consistent with the expected filtering properties of the **B** matrix in the analysis state and analysis increment equations. This suggests that the different representations of the analysis step (in the simulation of the error evolution) explain also an important part of the differences between the two methods.

Previous studies (Bouttier (1994) and McNally (2000)) suggest that the analysis error variances estimated by the NMC method are underestimated in the data-poor areas, and overestimated in regions where the observations are dense and have a high quality. For these reasons, in the Arpège 4D-Var, the NMC method is not used to provide the spatial variations of the background error variance. The examination of the local covariances, and the comparison between the analysis errors and the background errors estimated by the NMC method, are in agreement with these studies.

In contrast with this, the local ensemble background error variances capture some interesting features, such as the contrasts between data-rich and data-sparse regions, and also the areas of large atmospheric variability (the ITCZ, for instance). Moreover, the analysis dispersion provided by the ensemble approach appears to give some relevant information about the analysis error, for instance concerning the influence of data density.

The local correlation structures were also examined. For this purpose, an economical method was developed to estimate the local correlation length scales. The corresponding estimates provided by the ensemble method show that the length scale is largest in the tropics and smallest in the data rich regions, as it would be expected from other studies (e.g. Lindzen and Fox-

Rabinovitz (1989), Ingleby (2001), Bouttier (1994)). In addition, these estimates from the NMC method differ from the ensemble ones, mainly for temperature. In particular, the results suggest that the overestimation of the length scales by the NMC method is more pronounced over data rich areas, such as North America. These differences indicate that the expected data density effects on the correlation structures are not well represented by the NMC method.

Moreover, this local length scale diagnosis allows also to examine the correlation anisotropies. This information about heterogeneities and anisotropies can be useful to model spatially varying correlations in the **B** matrix, by wavelet techniques (Fisher (2003), Deckmyn and Berre (2005)) for instance.

Some impact studies were also performed with the Arpège 4D-Var. The use of the ensemble global covariances (compared with the NMC global covariances) had a general positive impact on the forecast quality. This positive impact was found to be enhanced when the ensemble local standard deviation maps (instead of the former operational ones) are used in the minimization and quality control stages. These results confirm the indications of the diagnostic studies: there is some relevant global and local information in the ensemble method, and this technique is in particular more appropriate than the NMC method. For the above reasons, the ensemble statistics became operational at Météo France in January 2004.

However, one could expect that the perfect model assumption causes an underestimation of the background error variances. Therefore, the impact of simulating the model error in the ensemble approach has also been studied. The model error is simulated by performing an ensemble of experiments using different resolutions and using different values for certain tuning parameters of some physical parametrizations.

The results indicate that the resolution is the most important contribution for small scale variables (e.g., vorticity and divergence). The perturbations of the physical parametrizations remain important in the tropics for vorticity and divergence, and also for temperature in general. This is more obvious in lower and middle troposphere in the extra-tropics and in upper levels over the tropics.

Moreover, the effect of the model error simulation in the ensemble approach is to increase the magnitude of the estimated background error. For large scale variables (temperature and surface pressure), the increase in magnitude of the background errors is associated to the enhancement of the synoptic scale contributions. On the other hand, for the other variables (vorticity, divergence, and humidity), the simulation of the model error appears to emphasize the relative contributions of the mesoscales to the background errors. This causes a reduction of the coupling between mass and wind variables, mainly due to the simulation of the model error related to resolution.

This reduction of the mass/wind coupling has a positive impact on the forecast quality of geopotential over the tropics and in some stratospheric extra-tropical areas. On the other hand, the reduction of mass/wind balance leads to a degradation of the forecasts of geopotential and wind in some tropospheric extra-tropical regions. This suggests that the mass/wind coupling is smaller in the tropics and in the stratosphere, and that the relaxation of this coupling as a function of latitude and height should be strengthened in the **B** matrix.

It was also found that there is a better agreement between the vertical profiles of the statistics of the innovations and those from the ensemble approach, when both the model error and observation errors are simulated.

The effect of neglecting the observation errors in the ensemble method has also been examined. For this purpose, the observation perturbations have been set to zero (the observations are supposed to be perfect). In other words, the ensemble simulations have been performed assuming that the background error is caused mainly by the model error. The resulting covariances have been compared with those in a perfect-model context, and with those when both the observation errors and the model errors are simulated.

The assumption that observations are perfect leads to a decrease of the background error

variance, mostly due to a reduction of the large scale contributions. This is consistent with the expected filtering properties of the analysis, which imply that the contribution of the observation errors to the analysis errors is particularly important in the large scales.

Furthermore, it was found that, globally, the magnitude of the background error estimated by the ensemble approach is smaller when it is assumed that the observations are perfect than when the model is assumed to be perfect. There are two possible interpretations for this result. The first one is that the uncertainties related to the observations have a larger contribution to the background error than the model error.

On the other hand, this result might also suggest that some important sources of model error have been underestimated. In the present work, the model error has been simulated using a simple approach of changing only certain tuning parameters from some parametrizations. In the future, it would be interesting to perform ensemble simulations using different (and independent) models, instead of different versions of the same model. These models should have different complete sets of parametrizations, e.g. for radiation, deep and shallow convection, vertical and horizontal diffusion and gravity wave drag. In addition, these models could have a different treatment of the orography.

The sensitivity of the ensemble estimates to the model resolution has also been studied. It was found that, when the model resolution increases, the magnitude of the background dispersion decreases in the large scales, and that it increases in the small scales. At first sight, the increase of small scale variance may be seen as a paradox. However, it is important to note that as the resolution of the model decreases, its variability at smaller scales will be smaller, because these scales are either not represented by the model, or dissipated by the horizontal diffusion. Thus, the amount of small scale dispersion is very dependent on the resolution and numerical diffusion. Moreover, the small scale structures that are observed but not represented at all by the model are usually considered and treated as a part of the observation representativeness errors, rather than a part of the background errors. More generally, on may wonder if the small scale variability that is dissipated by the numerical diffusion should also be considered and treated as a part of these representativeness errors. For the above reasons, the determination of the optimum amount of small scale variance, for specification of the $\mathbf{B}$ matrix in data assimilation, remains an open issue.

The ensemble simulations presented in this work have been done for a winter period. In the future, it would be also interesting to perform such experiments for other seasons. This would allow to examine the seasonal variations of e.g. the standard deviation maps (including for instance a detailed comparison with the results of the cycling algorithm that is described in Derber and Bouttier (1999)). More generally, the relevance of the ensemble local estimates suggests that there may be some interesting flow-dependent information to be extracted from the ensemble dispersion. In this perspective, it would be interesting to examine the potential of e.g. wavelets to diagnose and to specify these local flow-dependent signals.

Regarding the potential of the ensemble approach, it seems attractive to study its application also for limited area models. Thus, this technique has been applied to the Aladin model, which is coupled with the Arpège global model. In this context, the ensemble approach has been compared with two other error simulation techniques, regarding the representation of the analysis effect (Berre, Ştefănescu and Belo Pereira, 2006; appendix F). The evolution of the dispersion spectra and the evaluation of model differences (when applying the ensemble approach to the Aladin model) has also been studied (Ştefănescu, Berre and Belo Pereira, 2006; appendix G). The operational version of the Aladin-France 3D-Var is thus based on ensemble covariances, since its implementation in July 2005. The extraction of flow-dependent information from the ensemble of Aladin perturbed assimilation experiments is also a promising application.

# Appendix A

# Analysis Methods

## A.1 Optimal Interpolation

Optimal Interpolation (OI) is a method of data assimilation, which was operational in many numerical weather prediction centers during the 1970s, 80s and early 90s (McPherson *et al.*, 1979, Lorenc, 1981). OI is so called because it is a minimum variance estimator. This technique tries to find the **K** matrix which minimizes the analysis error variance.

In OI, the analysis field ($\mathbf{x}_a$) is obtained by replacing the matrix **K** (given by (1.14)) into the equation (1.13). Thus, the analysis estimate becomes:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\,\mathbf{B}\,\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{y} - \mathbf{H}\,\mathbf{x}_b) \tag{A.1}$$

It is interesting to notice that equation (A.1) is a vector generalization of the scalar expression (1.10).

It is also important to mention that, the following hypotheses are made in OI:

1) the analysis is an unbiased estimate of the true state;

2) the observation and background errors are uncorrelated;

3) the observation operator is linear, this approximation says that the variations of the observation operator $H$, in the vinicity of the background, are linear:

$$H(\mathbf{x}) = H(\mathbf{x}_b + \delta\mathbf{x}) \simeq H(\mathbf{x}_b) + \frac{\partial H(\mathbf{x}_b)}{\partial \mathbf{x}}\delta\mathbf{x} = H(\mathbf{x}_b) + \mathbf{H}\delta\mathbf{x} \tag{A.2}$$

where $\delta\mathbf{x} = \mathbf{x}_a - \mathbf{x}_b$ is the analysis increment.

It is also assumed that the analysis is optimal, i.e., that the analysis field represents the estimate closest to the true state according to the Root Mean Square Error (RMSE). However, it is important to remember that according to the estimation theory, the method can only be optimal if **B** and **R** matrices are known exactly. In practice, it is not possible to know them exactly, it is only possible to obtain an estimate of these matrices. For this reason, OI is also referred as Statistical Interpolation.

In OI, in order to determine the analysis field, it is necessary to invert the $\mathbf{HBH}^T + \mathbf{R}$ matrix (see equation A.1). However, the size of this matrix is $m \times m$, where $m$ is the size of observation vector and typically $m = 10^5$, so the inversion of this matrix is too expensive. Therefore, in order to reduce the matrix size, data selection is used. This means that it is assumed that each analysis point is only influenced by observations located in its closest vicinity. Hence, the observations used to perform the analysis at two neighboring points may be different. As a consequence, the analysis field will generally be spatially noisy.

## A.2 3D-Var

### A.2.1 Formulation and preconditioning

The basic aim of the Three-Dimensional Variational (3D-Var) Data Assimilation method is to find the model state (analysis, $\mathbf{x}_a$) which minimizes the cost function:

$$J = J_b + J_o = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(H(\mathbf{x}) - \mathbf{y})^T \mathbf{R}^{-1}(H(\mathbf{x}) - \mathbf{y}) \tag{A.3}$$

Here $J_b$ measures the distance of the model state $\mathbf{x}$ to a background model state $\mathbf{x}_b$ and $J_o$ measures the distance of $\mathbf{x}$ to the vector $\mathbf{y}$ of the observations.

The vector $\mathbf{x}_a$ which minimizes $J$ can be interpreted as the field that best fits simultaneously the background field and the observations.

When the cost function reaches its minimum, its derivative vanishes and the estimate $\mathbf{x}_a$ satisfies:

$$\nabla_{\mathbf{x}} J = 0 = \mathbf{B}^{-1}(\mathbf{x}_a - \mathbf{x}_b) + \mathbf{H}^T \mathbf{R}^{-1}(H(\mathbf{x}_a) - \mathbf{y}) \tag{A.4}$$

It is possible to prove that for linear observation operators, the solution of the 3D-Var method also satisfies the OI equation (A.1) (Lorenc, 1986). This shows that theoretically OI and 3D-Var are equivalent. However, in practice, the methodology of these two methods is rather different. 3D-Var avoids the inversion of the $\mathbf{HBH}^T + \mathbf{R}$ matrix by solving instead a global minimization problem, which is done iteratively using descent algorithms such as conjugate gradient method or quasi-Newton algorithms. Therefore, one of the major advantages of 3D-Var relatively to OI is that all observations over the Earth can be used simultaneously. In other words, in 3D-Var, data selection is not necessary. Moreover, the variational framework is more suited to the assimilation of observations indirectly related with the variables of the forecast model. For instance, the direct assimilation of radiance measurements is possible once a direct model and its adjoint are available (Eyre, 1989a, b). For these reasons, 3D-Var become operational in the major NWP centers in the 90s (Parrish and Derber, 1992; Rabier *et al.*, 1998, Courtier *et al.* and Gauthier *et al.*, 1999).

In 3D-Var, the accuracy of the analysis field is given by the equation (1.16) or (1.17), if $\mathbf{B}$ and $\mathbf{R}$ are exactly specified. Furthermore, it can be demonstrated that, in this case, the matrix $\mathbf{A}$ is equal to half of the inverse of the Hessian of $J$ (second order derivative of $J$).

Ideally, the analysis cost function should be specified in terms of fields which have the same resolution as the forecast model. However, in practice, $J$ is defined as a function of the analysis increments $\delta\mathbf{x} = \mathbf{x}_a - \mathbf{x}_b$, where $\delta\mathbf{x}$ has a lower resolution than the forecast model. This formulation is known as incremental formulation.

Applying the tangent linear approximation to the observation operator, the cost function for the incremental formulation is defined as follows

$$J = J_b + J_o = \frac{1}{2}\delta\mathbf{x}^T \mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}(\mathbf{H}\delta\mathbf{x} - \mathbf{d})^T \mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}) \tag{A.5}$$

where $\mathbf{d} = \mathbf{y} - H(\mathbf{x}_b)$ is the innovation vector.

In this case, the gradient of the cost function with respect to the increment $\delta\mathbf{x}$ is given by:

$$\nabla_{\delta\mathbf{x}} J = \mathbf{B}^{-1}\delta\mathbf{x} + \mathbf{H}^T \mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}) \tag{A.6}$$

However, in practice, the descent methods will not be applied to the previous equation, because it is possible to achieve a better preconditioning for the 3D-Var problem, and the convergence of a minimization algorithm is very sensitive to its preconditioning. An ideal preconditioning is obtained if the Hessian matrix is an identity matrix. One way of improving the preconditioning of the Hessian matrix is to define a matrix $\mathbf{L}$ and a new control variable $\chi$ such that

$$\mathbf{B} = \mathbf{L}\mathbf{L}^T \quad \Rightarrow \quad \mathbf{B}^{-1} = \mathbf{L}^{-T}\mathbf{L}^{-1} \tag{A.7}$$

and

$$\chi = \mathbf{L}^{-1}\delta\mathbf{x} \tag{A.8}$$

Therefore,

$$J_b = \frac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} = \frac{1}{2}\delta\mathbf{x}^T\mathbf{L}^{-T}\mathbf{L}^{-1}\delta\mathbf{x} = \frac{1}{2}\chi^T\chi \tag{A.9}$$

where $\chi$ is a transformed analysis increment.

Thus, in the space of $\chi$, the background error covariance matrix is equal to the identity matrix, and the cost function is expressed as:

$$J = \frac{1}{2}\chi^T\chi + \frac{1}{2}(\mathbf{HL}\chi - \mathbf{d})^T\mathbf{R}^{-1}(\mathbf{HL}\chi - \mathbf{d}) \tag{A.10}$$

and the gradient of the cost function with respect to the control variable $\chi$ is:

$$\nabla_\chi J = \chi + \mathbf{L}^T\mathbf{H}^\mathbf{T}\mathbf{R}^{-1}(\mathbf{HL}\chi - \mathbf{d}) \tag{A.11}$$

The problem defined by the previous equation is much better conditioned than the original minimization problem. For this reason, in practice, the minimization will be carried out in the space of $\chi$. Moreover, the initial point of the minimization is the background, so initially $\delta\mathbf{x} = \chi = 0$. It is interesting to notice that the preconditioned formulation of the 3D-Var method has the additional advantage to avoid the inversion of the **B** matrix.

The minimum of $J$ occurs when the norm of $\nabla_\chi J$ is zero. In practice, it is only possible to reduce the gradient norm to a small positive number. To reduce the computational cost of 3D-Var, the number of iterations performed during the minimization can be reduced arbitrarily, or requiring that the norm of the gradient of $J$ decreases by a predefined amount, which is an intrinsic measure of how much the analysis is close to its optimum.

At the end of minimization, the analysis increments are reconstructed in model space by $\delta\mathbf{x} = \mathbf{L}\chi$.

### A.2.2 The B matrix in 3D-Var

In order to illustrate the role of the **B** matrix in 3D-Var, suppose that there is a single observation of a model field (for instance, temperature) at one grid point, corresponding to the $j^{th}$ element of the model state vector. In this case, the observation operator is very simple and it is represented by the $1 \times m$ matrix whose $j^{th}$ element is equal to 1 and all other elements are zero:

$$\mathbf{H} = (0, 0, ..., 0, 1, 0, ..., 0) \tag{A.12}$$

It is important to remember that in 3D-Var method a cost function is minimized. Multiplying the equation (A.4) by **B** and rearranging, the analysis equation becomes

$$\mathbf{x}_a - \mathbf{x}_b = \mathbf{B}\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}_a) \tag{A.13}$$

In this simple example, $\mathbf{B}\mathbf{H}^T$ is equal to the $j^{th}$ column of **B**. Also, since there is one single observation, then $\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\,\mathbf{x}_a)$ is the scalar value $(y - x_a^j)/\sigma_o^2$, where $x_a^j$ is the analysed grid point value corresponding to the observation. Hence:

$$\mathbf{x}_a - \mathbf{x}_b = \left(\frac{(y - x_a^j)}{\sigma_o^2}\right)\begin{pmatrix} B_{1j} \\ B_{2j} \\ ... \\ B_{nj} \end{pmatrix} \tag{A.14}$$

Thus, the analysis increment is proportional to a column of the background error covariance matrix, **B**. In other words, the background covariance matrix controls in particular how the information is spread out from the single observation, to provide statistically consistent increments at the neighbouring grid points and levels of the model, and to ensure that observations of one model variable (e.g. temperature) produce dynamically consistent increments in the other variables (e.g. vorticity and divergence).

## A.3 4D-Var

### A.3.1 Formulation

In OI and 3D-Var methods, it is assumed that observations are available at the time when the analysis is performed. With such an approximation, it is not possible to take into account the temporal variability of observations within the analysis cycle. Depending on the type of observation, this can lead to an important loss of information, particularly for high frequency observations such as satellite and radar data. Therefore, the Four-Dimensional Variational (4D-Var) Data Assimilation was implemented operationally in the major NWP Centers during the late 90s or early 2000s, in order to overcome this problem (Rabier *et al.*, 2000; Mahfouf and Rabier, 2000).

The idea of 4D-Var is to find the initial state, $\mathbf{x}_0$, which produces a model trajectory (when integrated in time using the forecast model) that best fits the observations within the assimilation interval (known as the assimilation window) . Mathematically, in 4D-Var, the aim is to find the model state at the beginning of the assimilation window, $\mathbf{x}_0$, which minimizes the cost function:

$$J = J_b + J_o = \frac{1}{2}(\mathbf{x}(t_0) - \mathbf{x_b}(t_0))^T \mathbf{B}_0^{-1}(\mathbf{x}(t_0) - \mathbf{x_b}(t_0)) + \frac{1}{2}\sum_{i=0}^{N}(\mathbf{H}[\mathbf{x}_i] - \mathbf{y}_i)^T \mathbf{R}_i^{-1}(\mathbf{H}[\mathbf{x}_i] - \mathbf{y}_i) \quad \text{(A.15)}$$

where $N$ is the number of observation vectors $\mathbf{y}_i$ distributed within a time interval $[t_0, t_N]$ and $\mathbf{x}_i$ represents the model state at instant $t_i$, which is related to the initial state through the forecast model ($\mathbf{M}$):

$$\mathbf{x}_i = \mathbf{M}_i(\mathbf{x}(t_0)) = \mathbf{M}_i(\mathbf{x}_0) \quad \text{(A.16)}$$

The state $\mathbf{x}_0$ is called control variable because it is the variable with respect to which the minimization will be performed. In order to find the $\mathbf{x}_0$ which minimizes $J$ it is necessary to compute the gradient of $J_o$ and $J_b$. The gradient of the background component of the cost function is the same as in the 3D-Var.

The determination of $J_o$ gradient with respect to $\mathbf{x}_0$ is more complicated because of the strong constraint imposed by equation (A.16). Nevertheless, it is easy to compute the gradient of $J_o$ with respect to $\mathbf{x}_i$, which is given by

$$\nabla_{\mathbf{x}_i} J_o = H_i^T \mathbf{R}_i^{-1}(\mathbf{y}_i - H(\mathbf{x}_i)) \quad \text{(A.17)}$$

Thus, in order to determine the gradient of $J_o$ with respect to the control variable, it is necessary to establish the relationship between it and the gradient with respect to the state $\mathbf{x}_i$. For doing that is necessary to define the adjoint model. In NWP practice, for saving computation time, it is defined the Tangent Linear Model and its Adjoint.

A Tangent Linear Model (TLM) is obtained by linearizing the forecast model around its nonlinear trajectory, so that if a perturbation $\mathbf{dx}_{i-1}$ is introduced in the initial conditions, the final state is given by

$$\mathbf{x}_i + \mathbf{dx}_i = \mathbf{M}_{i-1}(\mathbf{x}_{i-1} + \mathbf{dx}_{i-1}) \simeq \mathbf{M}_{i-1}(\mathbf{x}_{i-1}) + \mathbf{L}_{i-1}\mathbf{dx}_{i-1} \quad \text{(A.18)}$$

Here $\mathbf{L}_{i-1}$ represents the Tangent Linear Model, which transforms an initial perturbation at $t_{i-1}$ to a final perturbation at $t_i$. Thus, the TLM equation is

$$\mathbf{dx}_i = \mathbf{L}_{i-1}\mathbf{dx}_{i-1} \quad \text{(A.19)}$$

so the TLM describes the evolution in time of perturbations around the initial state. The model $\mathbf{L}$ is called linear since it neglects the higher order terms associated to the nonlinear evolution of perturbations (tangent linear hypothesis).

The adjoint[2], $\mathbf{L}_i^*$, of a linear operator, $\mathbf{L}_i$, is defined as:

$$\langle \mathbf{L}_i x, z \rangle = \langle x, \mathbf{L}_i^* z \rangle \tag{A.20}$$

The first order variations in $J_o$ due to variations in the initial state (control variable) can be expressed as:

$$\delta J_o = < \nabla_{\mathbf{x}_0} J_o, \delta x_0 > \tag{A.21}$$

and they can also be written as:

$$\delta J_o = \sum_{i=0}^{N} < \nabla_{\mathbf{x}_i} J_o, \delta x_i > \quad \Leftrightarrow \quad \delta J_o = \sum_{i=0}^{N} < \nabla_{\mathbf{x}_i} J_o, \mathbf{L}_{i-1}\mathbf{L}_{i-2}...\mathbf{L}_0 \delta x_0 > \tag{A.22}$$

The equivalence in the last equation results from the tangent linear hypothesis and from the fact that the integration of TLM can be expressed as the product of intermediate forecast steps between the observations instants.

Applying the definition of the adjoint, it is possible to write the previous equation as

$$\delta J_o = \sum_{i=0}^{N} < \mathbf{L}_0^*\mathbf{L}_1^*...\mathbf{L}_{i-1}^* \nabla_{\mathbf{x}_i} J_o, \delta x_0 > = < \sum_{i=0}^{N} \mathbf{L}_0^*\mathbf{L}_1^*...\mathbf{L}_{i-1}^* \nabla_{\mathbf{x}_i} J_o, \delta x_0 > \tag{A.23}$$

Examining the equations (A.21) and (A.23), the following expression for calculation of the gradient of $J_o$ with respect to the control variable is obtained:

$$\nabla_{\mathbf{x}_0} J_o = \sum_{i=0}^{N} \mathbf{L}_0^*\mathbf{L}_1^*...\mathbf{L}_{i-1}^* \nabla_{\mathbf{x}_i} J_o \tag{A.24}$$

The last equation shows that the adjoint model propagates a perturbation backward in time, from instant $t_N$ to the initial time $t_0$.

Substituting the expression (A.17) in the previous equation and using a euclidean product, the gradient of $J_o$ with respect to the control variable is

$$\nabla_{\mathbf{x}_0} J_o = \sum_{i=0}^{N} \mathbf{L}_0^T\mathbf{L}_1^T...\mathbf{L}_{i-1}^T H_i^T \mathbf{R}_i^{-1}(\mathbf{y}_i - H(\mathbf{x}_i)) \tag{A.25}$$

Hence, the computation of $\nabla_{\mathbf{x}_i} J_o$ requires the computation of the increments $(\mathbf{y}_i - H(\mathbf{x}_i))$ at the observation times $t_i$ through a forward model integration, followed by its multiplication by $H_i^T \mathbf{R}_i^{-1}$, to obtain weighted increments. Finally, it requires the integration of these weighted observation increments backward in time to the initial time, using the adjoint model.

In summary, in 4D-Var, for each iteration $k$, the following steps are involved:

1. Integrate the nonlinear model with initial conditions, $\mathbf{x}_0^k$, from $t_0$ to $t_N$ and save the nonlinear trajectory in the assimilation window. This procedure is needed to define $\mathbf{L}_i^T$ and the observation increments $(\mathbf{d}_i = \mathbf{y}_i - H(\mathbf{x}_i))$.
2. Computation of the gradient of $J_b$, as in 3D-Var.
3. Computation of the gradient of $J_o$, using the expression (A.25), which includes the integration of the adjoint model.
4. Minimization of $J$ and modification of the control variable, defining $\mathbf{x}_0^{k+1}$.
5. All the previous steps are repeated until the minimum of the cost function is found.

Thus, in 4D-Var, each iteration of the minimization involves one integration of the nonlinear model and one integration of the adjoint model, which is 2 to 3 times more expensive than the

[2]the definition of the adjoint of a linear operator depends on the scalar product $< ., . >$ used; for instance, for the Euclidean product, the adjoint of a real matrix $\mathbf{L}$ is its transpose: $\mathbf{L}^* = \mathbf{L}^T$

nonlinear model. If 40 iterations are necessary to reach the minimum of $J$, then the cost of 4D-Var is 120 to 160 times larger than the cost of the forecast model integration. Thus, 4D-Var is very expensive. In practice, the 4D-Var problem is solved using the incremental formulation, in order to reduce the computational cost.

In the incremental formulation the cost function for 4D-Var is written as:

$$J = \frac{1}{2}(\delta\mathbf{x}(t_0))^T\mathbf{B}_0^{-1}(\delta\mathbf{x}(t_0)) + \frac{1}{2}\sum_{i=0}^{N}(\mathbf{H}_i\mathbf{L}_i\delta\mathbf{x}(t_0) - \mathbf{d}_i)^T\mathbf{R}_i^{-1}(\mathbf{H}_i\mathbf{L}_i\delta\mathbf{x}(t_0) - \mathbf{d}_i) \qquad \text{(A.26)}$$

where the tangent linear hypothesis was used to obtain

$$H_i(\mathbf{M}_i[\mathbf{x}(t_0)]) - \mathbf{y}_i = H_i(\mathbf{M}_i[\mathbf{x}_b + \delta\mathbf{x}(t_0)]) - \mathbf{y}_i \simeq H_i(\mathbf{M}_i(\mathbf{x}_b)) + \mathbf{H}_i\mathbf{L}_i\delta\mathbf{x}(t_0) - \mathbf{y}_i = \mathbf{H}_i\mathbf{L}_i\delta\mathbf{x}(t_0) - \mathbf{d}_i$$

In 4D-Var, the TLM can be obtained by linearizing the forecast model in the vinicity of a low resolution trajectory, and also by using either a package of simplified physics (Janisková *et al.*, 1999) or a more complex physics.

In operational 4D-Var, for the first iteration $\mathbf{x}_0 = \mathbf{x}_b$. In addition, the non-linear forecast model is used only once within a given minimization, for the first iteration. This allows to build the first trajectory, which propagates $\mathbf{x}_b$ within the assimilation window. For the next iterations, $\delta\mathbf{x}(t_0)$ are propagated using the TLM (necessary to compute the observation increments) and $\mathbf{x}_0$ is updated for $M$ iterations: $\mathbf{x}_0^{k+1} = \mathbf{x}_0^k + \delta\mathbf{x}_0^k$ and the analysis at instant $t_0$ is $\mathbf{x}_0^a = \mathbf{x}_0^M$.

The number of iterations, $M$, during the minimization of the gradient of $J$ are also reduced to diminish the computational cost of 4D-Var, as in the case of 3D-Var.

The Arpège 4D-Var uses an assimilation window of 6 hours centered around each synoptic time (06, 12, 18 or 00 UTC). This means that, for instance, the analysis, $\mathbf{x}_a$, at 12 UTC (which constitutes the initial condition of the forecast from 12UTC) doesn't correspond to the state $\mathbf{x}_0$, valid at the beginning of the assimilation window (9 UTC in this case), but to the 3h propagation of this state.

### A.3.2 The B matrix in 4D-Var

If the tangent linear approximation can be expressed as $\mathbf{x}_i = \mathbf{L}_i(\mathbf{x}(t_0))$, where $\mathbf{L}_i$ represents the TLM, then the 4D-Var cost function given by equation (A.15) can be rewritten as

$$J = \frac{1}{2}(\mathbf{x}(t_0) - \mathbf{x_b}(t_0))^T\mathbf{B}_0^{-1}(\mathbf{x}(t_0) - \mathbf{x_b}(t_0)) + \frac{1}{2}\sum_{i=0}^{N}(\mathbf{H}\mathbf{L}_i(\mathbf{x}_0) - \mathbf{y}_i)^T\mathbf{R}_i^{-1}(\mathbf{H}\mathbf{L}_i(\mathbf{x}_0) - \mathbf{y}_i) \quad \text{(A.27)}$$

As seen before, for the analysis the gradient of the cost function is zero:

$$\mathbf{B}^{-1}(\mathbf{x}_a(t_0) - \mathbf{x}_b(t_0)) + \mathbf{L}_i^T\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{H}\mathbf{L}_i(\mathbf{x}_a(t_0)) - \mathbf{y}(t_i)) = 0 \qquad \text{(A.28)}$$

where $\mathbf{L}_i^T$ represents the integration of the adjoint of TLM from time $t_i$ to $t_0$.

Since there is only one observation, the expression $\mathbf{R}^{-1}(\mathbf{H}\mathbf{L}_i(\mathbf{x}_a(t_0)) - \mathbf{y}(t_i))$ is equal to the scalar $(y - x_a^j(t_i))/\sigma_o^2$. Multiplying the previous equation by $\mathbf{B}$ and rearranging, it can be seen that, in 4D-Var, the analysis increment is proportional to a column of $\mathbf{B}\mathbf{L}_i^T$:

$$\mathbf{x}_a(t_0) - \mathbf{x}_b = \left(\frac{(y - x_a^j(t_i))}{\sigma_o^2}\right)\begin{pmatrix}(BL_i{}^T)_{1j}\\(BL_i{}^T)_{2j}\\\cdots\\(BL_i{}^T)_{nj}\end{pmatrix} \qquad \text{(A.29)}$$

Multiplying the equation (A.28) by $\mathbf{L_i B}$ and rearranging, the following equation is obtained:

$$\mathbf{L}_i(\mathbf{x}_a(t_0) - \mathbf{x}_b) = \left(\frac{(y - x_a^j(t_i))}{\sigma_o^2}\right) \begin{pmatrix} (L_i B {L_i}^T)_{1j} \\ (L_i B {L_i}^T)_{2j} \\ \cdots \\ (L_i B {L_i}^T)_{nj} \end{pmatrix} \tag{A.30}$$

and it is interesting to notice that the expression on the left hand side is:

$$\mathbf{L}_i(\mathbf{x}_a(t_0) - \mathbf{x}_b) = \mathbf{L}_i(\mathbf{x}_a(t_0)) - \mathbf{L}_i(\mathbf{x}_b) = \mathbf{x}_a(t_i) - \mathbf{L}_i(\mathbf{x}_b) \tag{A.31}$$

Since $\mathbf{B}$ is the background error covariance matrix at time $t_0$, then the matrix $\mathbf{L_i B L_i^T}$ is the covariance matrix for errors in a forecast from $t_0$ to $t_i$, with initial conditions equal to the background.

Under the tangent linear assumption and for a perfect model, the background errors at time $t_0$ and $t_i$ are related by

$$\mathbf{e}_b(t_i) = \mathbf{L}_i \mathbf{e}_b(t_0) \tag{A.32}$$

Thus, the background error covariance matrix at time $t_i$ is:

$$\overline{\mathbf{e}_b(t_i)(\mathbf{e}_b(t_i))^T} = \mathbf{L}_i \overline{\mathbf{e}_b(t_0)(\mathbf{e}_b(t_0))^T} \mathbf{L}_i^T = \mathbf{L_i B L_i^T} \tag{A.33}$$

Combining equations (A.30) and (A.31) it can be seen that the difference between the analysis and the forecast for time $t_i$ with initial conditions given by the background at time $t_0$ is proportional to a column of $\mathbf{L_i B L_i^T}$. In other words, the difference between the analysis trajectory and the background trajectory, at the time of the observations is proportional to the covariance matrix that describes the errors in the background trajectory at the time of the observation. This matrix is implicitly propagated or evolved by 4D-var using the dynamics of the TLM. Then, in 4D-Var, the analysis increment at the time of the observation is given by the column of the evolved covariance matrix (Thepaut *et al.*, 1993). Therefore, both the covariance matrix at the observation time and the analysis increments are flow-dependent. Nevertheless, the flow-dependent covariances which are used implicitly during the assimilation are not propagated to the next cycle.

# Appendix B

# The background cost function in ARPEGE 3D-Var and 4D-Var

## B.1 The square-root of the covariance matrix

This section describes the formulation of the **B** matrix that was used in the operational ARPEGE 4D-var until January 2005.

In this formulation, the balance relationships between mass and wind fields have an important role, because due to them the observations of the temperature field gives information about the wind field and vice-versa.

As in section A.2, the $J_b$ term can be written in the space of a new variable, $\chi$:

$$J_b = \frac{1}{2}\chi^T\chi \tag{B.1}$$

such that $\chi = \mathbf{L}^{-1}\delta\mathbf{x}$ and $\mathbf{L}$ verifies $\mathbf{L}\mathbf{L}^T = \mathbf{B}$.

In the operational practice, the initial point of the minimization is the background, so initially $\delta\mathbf{x} = \chi = 0$. The minimization is carried out in the space of $\chi$, as referred before in section A.2. At the end of the minimization, the analysis increments are reconstructed in the model space by $\delta\mathbf{x} = \mathbf{L}\chi$.

In this formulation, in order to calibrate the $\mathbf{L}$ operator, it is necessary to compute the balance operator $\mathcal{K}$ and the covariance matrix $\mathbf{B}_u$. The $\mathbf{B}$ matrix is not explicitly calculated and stored in memory, but rather implicitly by the definition of $\mathbf{L}$, which has the form:

$$\mathbf{L} = \mathcal{K}\,\mathbf{B}_u^{1/2} \tag{B.2}$$

where $\mathcal{K}$ is a balance operator that transforms the variables $(\zeta, \eta_u, (T, P_{surf})_u, q)$ of the control vector, to the model variables $(\zeta, \eta, (T, P_{surf}), q)$, in the following way:

$$\begin{array}{rcl} \zeta & = & \zeta \\ \eta & = & \mathbf{M}\zeta + \eta_u \\ (T, P_{surf}) & = & \mathbf{N}\zeta + \mathbf{P}\eta_u + (T, P_{surf})_u \\ q & = & q \end{array}$$

where $P_{surf}$ is the surface pressure.

## B.2 The balance operator $\mathcal{K}$

In the matrix from, according to Derber and Bouttier (1999), the balance operator $\mathcal{K}$ is given by

$$\mathcal{K} = \begin{pmatrix} I & 0 & 0 & 0 \\ \mathbf{M} & I & 0 & 0 \\ \mathbf{N} & \mathbf{P} & I & 0 \\ 0 & 0 & 0 & I \end{pmatrix} \tag{B.3}$$

The **M**, **N** and **P** balance operators have all the same structure: block-diagonal, with one full vertical matrix per spectral component. The vertical matrices depend only on the wave number $n$. **M** and **N** are both the product of a so-called horizontal balance operator $\mathcal{H}$ by vertical balance operators $M$, $N$:

$$\mathbf{M} = M\mathcal{H} \qquad \mathbf{N} = N\mathcal{H}$$

The $\mathcal{H}$ operator is a block-diagonal matrix of horizontal operators transforming the spectral coefficients of vorticity, independently at each level, into an intermediate variable $P_b$, which is a linearized mass variable defined as:

$$P_b(m,n) = \beta_1(n,m)\zeta(n,m+1) + \beta_2(n,m)\zeta(n,m-1) \tag{B.4}$$

The coupling between $P_b$ and $\zeta$ is justified by the geostrophic equilibrium, which on the sphere can be expressed by

$$\Phi_b = \Delta^{-1}\,\mathrm{div}(f\,\Delta^{-1}\zeta)$$

where $\Phi_b$ is the geostrophically balanced part of the geopotential, $f$ is the Coriolis parameter and $\Delta$ refers to the Laplacian operator.

It is important to mention that in $J_b$, the balance expressed by the equation (B.3) and (B.4) depends on latitude and on model levels. Obviously the geostrophic coupling is weaker in the lower troposphere and in tropics. This coupling is also weaker for the small horizontal and vertical scales of the fields than for the large scales.

The horizontal balance coefficients $(\beta_1, \beta_2)$ of $\mathcal{H}$ are computed by a linear regression between the errors in vorticity and in the linearized total mass $P_{tot}$, assuming the functional relationship defined by the equation $(B.4)$ with $P_{tot} = P_b$. $P_{tot}$ is calculated using the linearized hydrostatic relationship at level $l$:

$$P_{tot}(l) = \sum_{i=nlev}^{l} RT_i\,\Delta\ln p_i + RT_{ref}\ln P_{surf} \tag{B.5}$$

where $nlev$ is the number of model levels and $p_i$ refers to the pressure at the interfaces of the model layers. $T_{ref}$ is a reference temperature and the linear regression is only slightly sensitive to its choice.

The $M$ vertical operator is computed for each wave number $n$ by a linear regression between the spectral vertical profiles of balanced mass, $[P_b]_n^m$ and divergence, $[\eta]_n^m$, according to:

$$[\eta]_n^m = M(n)[P_b]_n^m + [\eta_u]_n^m \tag{B.6}$$

where $\eta_u$ is the residual of the regression and is also named the unbalanced part of divergence.

The coupling between $P_b$ and $\eta$ corresponds to the Ekman pumping, which causes winds to deviate from the geostrophic balance due to the presence of surface friction. This creates an inflow near the surface towards the center in extra-tropical low-pressure systems and a compensating upper-tropospheric outflow. This establishes a partial coupling between the divergent wind component and vorticity, which is much stronger in the middle latitudes, near the surface and in high troposphere.

The $N$ and $P$ operators are computed for each wave number $n$ exactly like $M$, except that now the linear regression is given by:

$$[(T, p_{surf})]_n^m = N_n[P_b]_n^m + P_n[\eta_u]_n^m + [(T, P_{surf})_u]_n^m \tag{B.7}$$

where $(T, P_{surf})_u$ is the residual of the regression. It is also called the unbalanced part of $(T, P_{surf})$.

## B.3 The covariance matrix $\mathbf{B}_u$

In equation (B.2) the operator $\mathbf{B}_u^{1/2}$ is a symmetric square root of $\mathbf{B}_u$. The background error covariance matrix $\mathbf{B}_u$ is assumed to be diagonal, with no correlation between the parameters:

$$\mathbf{B}_u = \begin{pmatrix} C_\zeta & 0 & 0 & 0 \\ 0 & C_{\eta_u} & 0 & 0 \\ 0 & 0 & C_{(T,P_{surf})_u} & 0 \\ 0 & 0 & 0 & C_q \end{pmatrix} \tag{B.8}$$

In the operational practice, the vertical autocovariances ($C$) of $\zeta, \eta_u, (T, P_{surf})_u, q$ are computed for each wave number $n$, using some forecast differences, provided by the NMC method. This method is described in chapter 2.

Since there are $2n+1$ wave numbers $m$ for each $n$ and for each forecast difference, it is necessary to have as many linearly independent differences as model levels, in order to ensure that the autocovariances are positive definite at the very large scales. Moreover, it is advisable to use several times more forecast differences than model levels in order to reduce the sampling noise at large scales, which is important for the performance of the assimilation/forecast system (Derber and Boutier, 1999).

Additionally, the variance spectra are slightly modified to ensure that the horizontal error correlations are compactly supported (they are set to zero beyond 6000 km). This operation removes the residual sampling noise in the error covariances.

The $\zeta, \eta_u, (T, P_{surf})_u$ vertical profiles of total standard deviation were estimated by the NMC method. Moreover, these vertical profiles were rescaled by an arbitrary factor ($R_{NMC}$) of 0.9, in order to account for the mismatch between the amplitudes of 12/36-hour forecast differences and of 6 hour forecast errors (Derber and Bouttier, 1999).

In this formulation, vorticity plays an important role, as its statistics determine the main part of the wind covariances, and also the main part of the temperature and surface pressure covariances. The vorticity correlations do not vary horizontally, but the vorticity standard deviations are allowed to vary geographically.

In the ECMWF 4D-Var, a cycling algorithm is used to determine some space and time variations of these vorticity standard deviations, as described in the appendix of Derber and Bouttier (1999). Since the average vertical profile of the standard deviation is provided by the NMC method, only the normalized values of vorticity standard deviations given by the cycling technique are used. In other words, the cycling algorithm is used only to modulate horizontally the average vertical profile of the standard deviations (provided by the NMC). This cycling algorithm is not implemented in the operational Arpège 4D-Var (at Météo France). Instead, the ECMWF maps that were produced for the 20th of October 1998 were used. As it is illustrated in section 3.4.1, these maps allow to represent e.g. some typical data density contrasts over the globe. Finally, the three-dimensional pattern of background error variance of $\zeta$ is obtained by multiplying the normalized background error variances by the global vertical profile of variances times $R^2_{NMC}$.

Moreover, some randomization techniques (Andersson and Fisher 1998) are applied to deduce some maps of standard deviations for wind, temperature and geopotential. These latter maps are used in the quality control of the Arpège 4D-Var, when a first-guess check is performed to reject observations whose values are too different from the background values.

The background error standard deviations for $q$ are specified in gridpoint space, in order to allow geographical variations, using an empirical function of temperature and specific humidity (Rabier *et al.*, 1998).

## B.4 The implied covariance matrix

It is important to notice that assuming that the unbalanced variables are uncorrelated is not too restrictive because the form of the $\mathcal{K}$ balance operator introduces significant correlations between the meteorological variables. The multiplication of $\mathbf{B}_u$ by $\mathcal{K}$ allows to write explicitly the implicit background error covariance matrix in terms of the meteorological variables $\zeta$, $\eta$, $(T, P_{surf})$ and $q$:

$$\mathbf{B} = \mathcal{K}\mathbf{B}_u\mathcal{K}^T = \begin{pmatrix} C_\zeta & C_\zeta\mathbf{M}^T & C_\zeta\mathbf{N}^T & 0 \\ \mathbf{M}C_\zeta & \mathbf{M}C_\zeta\mathbf{M}^T + C_{\eta_u} & \mathbf{M}C_\zeta\mathbf{N}^T + C_{\eta_u}\mathbf{P}^T & 0 \\ \mathbf{N}C_\zeta & \mathbf{N}C_\zeta\mathbf{M}^T + \mathbf{P}C_{\eta_u} & \mathbf{N}C_\zeta\mathbf{N}^T + \mathbf{P}C_{\eta_u}\mathbf{P}^T + C_{(T,P_{surf})_u} & 0 \\ 0 & 0 & 0 & C_q \end{pmatrix} \quad \text{(B.9)}$$

From last equation it is possible to see that only the specific humidity analysis is independent from the other variables. Moreover, the structure of the balance operator defines the form of the multivariate correlations between the meteorological variables. It is also important to notice that the vertical correlations in $C_\zeta$, the unbalanced divergence covariances $C_{\eta_u}$, and the unbalanced temperature covariances $C_{(T,P_{surf})_u}$ are both homogeneous and isotropic. Nevertheless, the terms involving the horizontal balance operator have a dependency on latitude.

As seen before, the structure of the horizontal operator $\mathcal{H}$ is related to the geostrophic balance, which is itself closely approximated by the equation $P_b = f\Delta^{-1}\zeta$, when the $f$-plane approximation is used. Thus, on a narrow band around a given extratropical latitude, $\varphi$, the covariance matrix of the balanced geopotential $C_{P_b} = \mathcal{H}C_\zeta\mathcal{H}^T$ can be approximated by a function of the Coriolis parameter and of the covariance of the vorticity background error according to:

$$C_{P_b} \simeq \hat{C}_{P_b}(\varphi) = f^2\Delta^{-2}C_\zeta$$

Substituting the expression $\mathcal{H}C_\zeta\mathcal{H}^T \simeq \hat{C}_{P_b}(\varphi)$ into equation (B.9), the effective $\mathbf{B}$ matrix in extratropical regions can be written as:

$$\mathbf{B}(\varphi) = \begin{pmatrix} C_\zeta & \hat{C}_{P_b}M^T & \hat{C}_{P_b}N^T & 0 \\ M\hat{C}_{P_b} & M\hat{C}_{P_b}M^T + C_{\eta_u} & M\hat{C}_{P_b}\mathbf{N}^T + C_{\eta_u}\mathbf{P}^T & 0 \\ N\hat{C}_{P_b} & M\hat{C}_{P_b}M^T + \mathbf{P}C_{\eta_u} & N\hat{C}_{P_b}N^T + \mathbf{P}C_{\eta_u}\mathbf{P}^T + C_{(T,P_{surf})_u} & 0 \\ 0 & 0 & 0 & C_q \end{pmatrix} \quad \text{(B.10)}$$

It is important to notice that the vertical correlations of $\zeta$ are independent of the latitude. However, in the matrix $\mathbf{B}(\varphi)$ there are two latitude dependent vertical autocovariances: the "vorticity-balanced divergence", $M\hat{C}_{P_b}M^T$, and the "vorticity-balanced $(T, p_{surf})$", $N\hat{C}_{P_b}N^T$. Moreover, all the terms that depend on $\hat{C}_{P_b}$ are modulated by the square of the sine of latitude according to $\hat{C}_{P_b}(\varphi) = f^2\Delta^{-2}C_\zeta$ (Derber and Bouttier, 1999).

In tropical regions, $f$ is very small, so $C_{P_b}$ is negligible when compared to the other covariance matrices and then the effective $\mathbf{B}$ matrix is

$$\mathbf{B} = \begin{pmatrix} C_\zeta & 0 & 0 & 0 \\ 0 & C_{\eta_u} & C_{\eta_u}\mathbf{P}^T & 0 \\ 0 & \mathbf{P}C_{\eta_u} & \mathbf{P}C_{\eta_u}\mathbf{P}^T + C_{(T,P_{surf})_u} & 0 \\ 0 & 0 & 0 & C_q \end{pmatrix} \quad \text{(B.11)}$$

In the tropics, there is a coupling between temperature and divergence, which is stronger near the ground and near the tropopause. For instance, an increase in temperature (or a decrease in $P_{surf}$) is associated with weak convergence near the ground and weak divergence near the tropopause.

In summary, when the gridpoint variances do not vary horizontally, the autocovariance model is non-separable, homogeneous and isotropic in gridpoint space. Thus, when the variances are

horizontally constant, the vertical correlation structures do not depend on geographical location, but they depend on the horizontal scale (Derber and Bouttier, 1999). Large horizontal scales have deeper vertical correlations than small horizontal scales. However, the vertical correlations of temperature and divergence vary also spatially, depending on latitude. The shape of the horizontal correlations is determined by the covariance spectra. Furthermore, the horizontal correlations are assumed to be spatially homogeneous and isotropic.

## B.5 The explained variance

The balance operator $\mathcal{K}$ is calibrated using a linear regression principle. According to Derber and Bouttier (1999), the performance of this operator can be quantified by diagnosing the explained variance ratio $V_{exp}$ from the meteorological fields, which is determined as:

$$V_{exp} = 1 - \frac{\sigma_{un}^2}{\sigma_t^2} \tag{B.12}$$

where $\sigma_t^2 = \sigma_{un}^2 + \sigma_{bal}^2$ is the total variance of the fields, $\sigma_{un}^2$ and $\sigma_{bal}^2$ are respectively the variances of the unbalanced and of the balanced components of the fields.

# Appendix C

# Local length scale and anisotropy

## C.1 An economical estimation of the local length scale

The background error variance of each variable and of its derivative can be used to estimate the horizontal correlation length scale. For the sake of simplicity, the derivations are presented for the unidimensional case. For any error $e$, the covariance of the error derivative $\frac{\partial e}{\partial x}$ between two points can be expressed as:

$$\left\langle \frac{\partial e_1}{\partial x_1} \frac{\partial e_2}{\partial x_2} \right\rangle = \frac{\partial^2}{\partial x_1 \partial x_2} (\langle e_1 e_2 \rangle) \tag{C.1}$$

Here $e_1 = e(x_1)$ and $e_2 = e(x_2)$, where $x_1$ and $x_2$ refer to the coordinates of the two points being correlated.

Furthermore, the background error covariance can be defined as a function of its autocorrelation function ($\rho$), and of its variances $\sigma(e_1)$ and $\sigma(e_2)$ at points $x_1$ and $x_2$, respectively:

$$\langle e_1 e_2 \rangle = \sigma(e_1)\sigma(e_2)\,\rho \tag{C.2}$$

Thus, the second derivative of $\langle e_1 e_2 \rangle$ is

$$\begin{aligned} \frac{\partial^2}{\partial x_1 \partial x_2} (\langle e_1 e_2 \rangle) = \frac{d\sigma(e_1)}{dx_1} \frac{d\sigma(e_2)}{dx_2} \rho + \sigma(e_1)\sigma(e_2) \frac{\partial^2 \rho}{\partial x_1 \partial x_2} + \\ + \sigma(e_1) \frac{d\sigma(e_2)}{dx_2} \frac{\partial \rho}{\partial x_1} + \sigma(e_2) \frac{d\sigma(e_1)}{dx_1} \frac{\partial \rho}{\partial x_2} \end{aligned} \tag{C.3}$$

Defining $\tilde{x} = x_2 - x_1$ allows to write that $\frac{\partial}{\partial x_1} = -\frac{d}{d\tilde{x}}$ and $\frac{\partial}{\partial x_2} = \frac{d}{d\tilde{x}}$. In addition, the first derivative of $\rho$ (it is supposed that $\rho$ is derivable) can be assumed to be equal to zero at $\tilde{x} = 0$. This assumption looks reasonable, knowing that the value of $\rho$ decreases when the separation distance increases from the origin ($\tilde{x} = 0$), and assuming that $\rho$ is continuous near its origin.

Replacing equation (C.1) into equation (C.3), and considering the case $\tilde{x} = 0$, provides the following expression for the local variance of $\frac{\partial e}{\partial x}$ (after replacing the notation $\tilde{x}$ by $x$):

$$\sigma^2\left(\frac{\partial e}{\partial x}\right) = \left(\frac{d\sigma(e)}{dx}\right)^2 - \sigma^2(e)\, \frac{d^2\rho}{dx^2}\bigg|_{x=0}$$

Finally, using the definition of the length scale given by Daley (1991), for the unidimensional case

$$L^2 = -\frac{\rho(x)}{d^2\rho/dx^2}\bigg|_{x=0} \tag{C.4}$$

we obtain the correlation length scale for the background error of any variable $e$, as a function of the background error variances of $e$ and $\frac{\partial e}{\partial x}$:

$$L_x^e = \sqrt{\frac{\sigma^2(e)}{\sigma^2(\frac{\partial e}{\partial x}) - \left(\frac{d\sigma(e)}{dx}\right)^2}} \tag{C.5}$$

It is important to emphasize that this estimate is only valid when using the length scale definition from Daley (1991).

It is interesting to note that due to the Helmholtz's theorem, for streamfunction ($\psi$) the correlation length scale can be written as a function of the background error variances of $\psi$ and of the rotational part of the meridional wind ($v^\psi = \frac{\partial \psi}{\partial x}$):

$$L_x^\psi = \sqrt{\frac{\sigma^2(\psi)}{\sigma^2(v^\psi) - \left(\frac{d\sigma(\psi)}{dx}\right)^2}}. \tag{C.6}$$

## C.2 Anisotropy: two simple examples

In order to illustrate the behaviour of the inertia matrix $\mathcal{N}$ of the correlation function $\rho$ (see equation 3.2), it is useful to consider two simple examples. Let us consider the following matrices:

$$\mathcal{N} = \mathcal{N}_m = \begin{pmatrix} 4 & -1 \\ -1 & 2 \end{pmatrix} \tag{C.7}$$

and

$$\mathcal{N} = \mathcal{N}_t = \begin{pmatrix} 2 & -1 \\ -1 & 4 \end{pmatrix} \tag{C.8}$$

Thus, the tilting term is the same for $\mathcal{N}_m$ and $\mathcal{N}_t$. However, for $\mathcal{N}_m$, $L_y^2 = 2\,L_x^2$ and for $\mathcal{N}_t$, $L_x^2 = 2\,L_y^2$.

These matrices have two eigenvectors and eigenvalues. The two eigenvalues are the same for $\mathcal{N}_t$ and $\mathcal{N}_m$. They are obtained by setting the determinant of $\mathcal{N} - \lambda \mathbf{I}$ equal to zero, which gives:

$$\lambda^2 - 6\lambda + 7 = 0$$

Thus, the eigenvalues are $\lambda_1 = 4.42$ and $\lambda_2 = 1.59$. Hence, the magnitude of the anisotropy of the correlation functions, given by the oblateness ($\mathcal{O}$) of $\mathcal{N}_t$ and $\mathcal{N}_m$ is: $\mathcal{O} = 1 - \lambda_2/\lambda_1 = 0.64$.

The two eigenvectors are found by substituting each of the eigenvalues into the matrix $\mathcal{N} - \lambda \mathbf{I}$.

Therefore, the eigenvector of the matrix $\mathcal{N}_m$ associated to the eigenvalue $\lambda_2$ is $\vec{X} = \vec{i_x} + 2.42\,\vec{i_y}$ and the normalized eigenvector is $\vec{X_n} = 0.38\,\vec{i_x} + 0.92\,\vec{i_y}$.

In this case, the main anisotropy axis, which identifies the direction of the largest elongation of the correlation function, is given by the following vector:

$$\vec{X}_{an}^m = 0.64\,\vec{X_n} = 0.24\,\vec{i_x} + 0.59\,\vec{i_y}.$$

For the matrix $\mathcal{N}_t$ the eigenvector associated to the eigenvalue $\lambda_2$ is $\vec{X} = \vec{i_x} + 0.42\,\vec{i_y}$ and the normalized eigenvector is $\vec{X_n} = 0.93\,\vec{i_x} + 0.38\,\vec{i_y}$. Therefore, the anisotropy vector is:

$$\vec{X}_{an}^t = 0.64\,\vec{X_n} = 0.60\,\vec{i_x} + 0.24\,\vec{i_y}.$$

The shapes of the two correlation functions associated to $\mathcal{N}_t$ and $\mathcal{N}_m$ are plotted in figure C.1.

In this figure it is visible that the correlation function associated to $\mathcal{N}_m$ is mainly elongated in the SW-NW direction, while for $\mathcal{N}_t$, $\rho$ is elongated in the WSW-ENE direction.

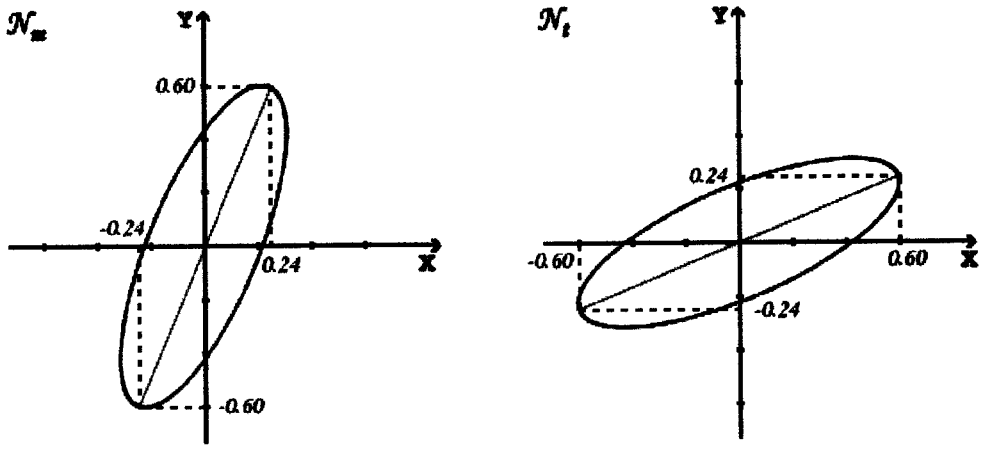


Figure C.1: Shape of the correlation function associated to $\mathcal{N}_m$ (left panel), and to $\mathcal{N}_t$ (right panel).

# Appendix D

# Innovations of temperature

This appendix presents the vertical profile of the RMSE computed for the innovation vector of temperature, for certain regions (see table 3.1). These computations were done for one month period between 21 of February to 21 of March 2002. This period is included in the period used for the ensemble simulations (4/2/2002 to 24/3/2002).

The innovation vector, $\mathbf{d} = \mathbf{y} - \mathbf{H}(\mathbf{x}_b)$, can be written as

$$\mathbf{d} = \mathbf{y} - \mathbf{x}_b = (\mathbf{y} - \hat{\mathbf{x}}) - (\mathbf{x}_b - \hat{\mathbf{x}}) = \mathbf{e}_o - \mathbf{e}_b$$

if the model grid coincides with the observation location, i.e., $\mathbf{H} = \mathbf{I}$.

In this case, if the observation and background errors are not correlated, the covariance of the innovation vector can be written as:

$$\langle (\mathbf{y} - \mathbf{x}_b)(\mathbf{y} - \mathbf{x}_b)^T \rangle = \langle (\mathbf{e}_o)(\mathbf{e}_o)^T \rangle + \langle (\mathbf{e}_b)(\mathbf{e}_b)^T \rangle$$

Thus, the variance of the innovation, $\sigma_d^2$, is the sum of the background and observation error variances: $\sigma_d^2 = \sigma_b^2 + \sigma_o^2$.

In practice, $\mathbf{H} \neq \mathbf{I}$. Therefore, $\sigma_d^2$ also contains the contribution of the representativeness error, which corresponds to small scales that are observed but not represented by the model (and which are a part of the observation errors, rather than a part of the background errors).

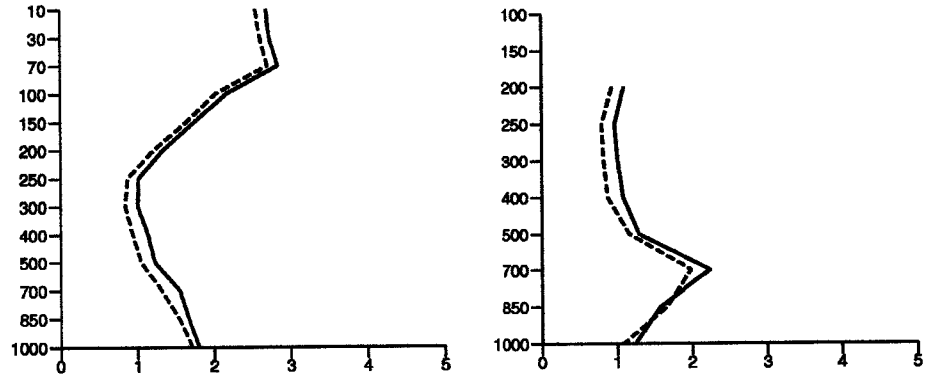


Figure D.1: RMSE of the innovation vector (solid line) and of the observation-minus-analysis differences (dashed line) for temperature (K) over the TROPIQ region, computed with observations from radiosondes (left panel) and from AIREP (right panel).

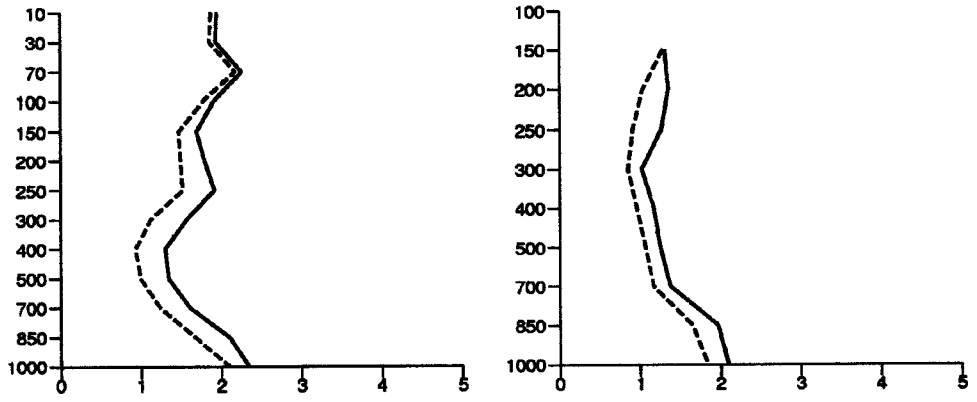


Figure D.2: RMSE of the innovation vector (solid line) and of the observation-minus-analysis differences (dashed line) for temperature (K) over the SUD20 region, computed with observations from radiosondes (left panel) and from AIREP (right panel).

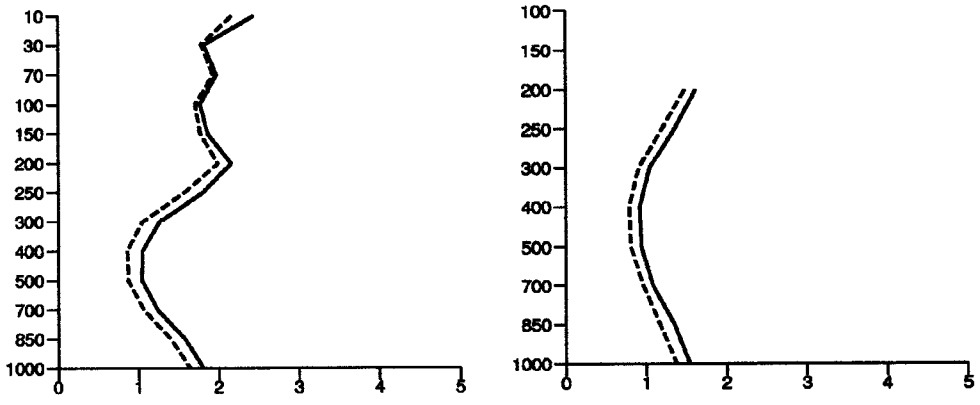


Figure D.3: RMSE of the innovation vector (solid line) and of the observation-minus-analysis differences (dashed line) for temperature (K) over the EUROPE region, computed with observations from radiosondes (left panel) and from AIREP (right panel).

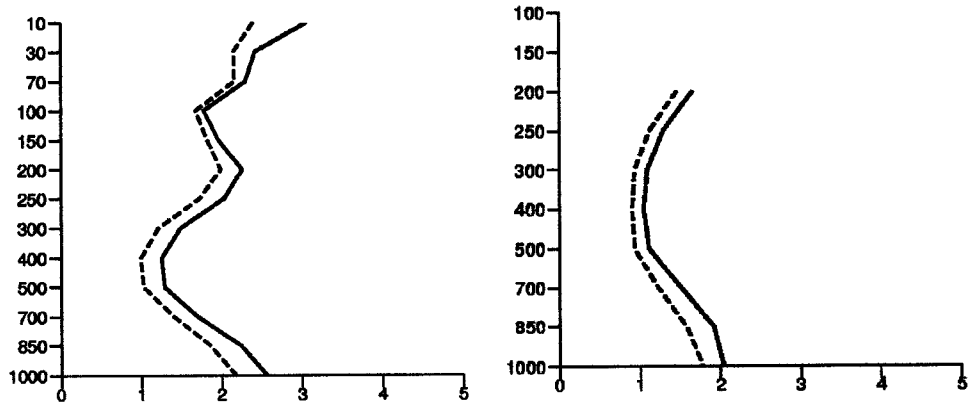


Figure D.4: RMSE of the innovation vector (solid line) and of the observation-minus-analysis differences (dashed line) for temperature (K) over the AMERNOR region, computed with observations from radiosondes (left panel) and from AIREP (right panel).

# Appendix E

# Changes in resolution, horizontal diffusion and physical parametrizations

In order to simulate the model error in the ensemble approach, some tuning parameters from certain parametrizations have been perturbed. The modifications for vertical diffusion, cloudiness and deep convection were chosen among some changes which became operational for Arpège in April 2003. The T449 resolution and its specific coefficients for horizontal diffusion correspond to some experimental high resolution trials, which were also conducted during this period for Arpège. A brief description of these parametrizations is introduced here. In addition, the changes in the tunable parameters of these parametrizations are presented.

## E.1 Summary of the model changes

The model changes are the following, by decreasing order of expected importance:

- Increase of resolution and horizontal diffusion:
  - o The coefficients of horizontal diffusion have been adjusted in accordance with the increase of horizontal resolution (T449 instead of T299). The main expected effect is that small scale structures are less dissipated than in the T299 integration.
- Turbulent fluxes:
  - o Some vertical diffusion parameters have been changed, with the effect that turbulent fluxes are smaller in stable cases, which reduces erosion of low layer inversions.
- Cloudiness scheme ; the spatial extension and the amounts of cloud condensates have been reduced, by introducing the following modifications:
  - o A parameter that controls the triggering of stratiform clouds has been changed.
  - o A parameter that controls the shallow convection effects on cloudiness (at the top of the Planetary Boundary Layer (PBL)) has been modified.
  - o Less importantly, a parameter that controls the effect of convective precipitation on the convective condensate has been changed.

- Deep convection:

  o The contribution of the large scale precipitation flux to the available humidity (for deep convection) has been removed.

  o The compensation of turbulent fluxes by convective fluxes has been changed to reduce the cyclogenetic tendencies.

  o The modulation factor of moisture convergence has been increased, which means that there is more available humidity for deep convection.

## E.2 Horizontal diffusion scheme

In the absence of horizontal diffusion, enstrophy will accumulate at the smallest scales of the NWP model, instead of being transferred to the very small scales, where it would be dissipated. It is thus necessary to apply "some" horizontal diffusion. At the model top, horizontal diffusion is also useful to filter waves reflected spuriously at the rigid top boundary of the model. The contribution of the horizontal diffusion to the tendency of any variable $\psi$ is the following one:

$$\left(\frac{\partial \psi}{\partial t}\right)_{horiz\,diffus} = -\mathcal{K}_\psi \, m_p \, \nabla'^r \, \psi$$

where $r$ is the power of the horizontal diffusion scheme ($r = 4$ in Arpège), and $\nabla'^r$ is the reduced gradient operator. The diffusion coefficient $\mathcal{K}_\psi$ is given by:

$$\mathcal{K}_\psi = exp(-0.5\pi\, i\, r)\left(\sqrt{\frac{N_s(N_s+1)}{r_a^2}}\right)^{-r} \Omega\, h_\psi\, g$$

where $\Omega$ is the angular velocity of the Earth rotation and $h_\psi$ is a constant coefficient for each prognostic variable. The values of the coefficients $\frac{1}{\Omega h_\psi}$ for divergence (HDIRDIV), for vorticity (HDIRVOR), for temperature (HDIRT) and for specific humidity (HDIRQ) used in experiments are presented in table E.1.

Table E.1: Values of the constant coefficients used in the horizontal diffusion scheme.

| Parameters | EXP T299 | HD(T449) |
|---|---|---|
| HDIRDIV | 1085 | 362 |
| HDIRVOR | 5425 | 1810 |
| HDIRT | 5425 | 1810 |
| HDIRQ | 5425 | 1810 |

## E.3 Turbulent fluxes

According to the K-theory it is assumed that the turbulent eddies behaves similarly to molecular diffusion. Therefore, the turbulent fluxes $F'_\psi$ for any conservative variable $\psi$ can be considered to be proportional to the local gradient of the mean field:

$$\frac{\partial \psi}{\partial t} = \frac{1}{\rho}\frac{\partial F_\psi}{\partial z} \quad and \quad F_\psi = \rho K_\psi \left(\frac{\partial \overline{\psi}}{\partial z}\right) \tag{E.1}$$

where $K_\psi$ is the eddy exchange coefficient. In the Arpège model, $\psi$ represents $u, v, s, q$ ($s = c_p T' + g z$ is the dry static energy).

### E.3.1 Surface layer

In the surface layer the turbulent fluxes are parametrized following the Monin-Obukhov's similarity theory. Thus, they can be written as

$$F_\psi = \rho C_\psi |V_N| (\psi_N - \psi_S) \tag{E.2}$$

where $C_\psi$ represents the surface turbulent exchange coefficient for the variable $\psi$. The indexes $N$ and $S$ refer to the lowest model level and the surface level, respectively.

The exchange coefficients for momentum are given by

$$C_M = \begin{cases} \left( \frac{1}{1+2b\left(\frac{R_i}{\sqrt{1+d\,R_i}}\right)} \right) C_M^n & \text{if } R_i > 0, \text{ stable case} \\ \left( 1 - \frac{2b\,R_i}{1+3bcC_M^n\left(\sqrt{\frac{z+z_o}{z_o}(-R_i)}\right)} \right) C_M^n & \text{if } R_i < 0, \text{ unstable case} \end{cases}$$

and for energy (dry static energy and specific humidity) are given by

$$C_H = \begin{cases} \left( \frac{1}{1+3b\,R_i\left(\sqrt{1+d\,R_i}\right)} \right) C_H^n & \text{if } R_i > 0, \text{ stable case} \\ \left( 1 - \frac{3b\,R_i}{1+3bcC_H^n\left(\sqrt{\frac{z+z_{oh}}{z_{oh}}(-R_i)}\right)} \right) C_H^n & \text{if } R_i < 0, \text{ unstable case} \end{cases}$$

where $b$, $c$ and $d$ are tunable parameters, $z_o$ is the surface dynamical roughness length, $z_{oh}$ is the surface thermal roughness length, and $R_i$ is the Richardson number, which characterizes the stability of the atmosphere and is given by

$$R_i = \frac{\frac{g}{\theta}\frac{\partial \theta}{\partial z}}{\left|\frac{\partial V}{\partial z}\right|^2} \sim \frac{\frac{g}{c_p T}\frac{\partial s}{\partial z}}{\left|\frac{\partial V}{\partial z}\right|^2}$$

$C_M^n$ and $C_H^n$ are respectively the exchange coefficients in neutral conditions for momentum and energy, which are computed as function of $z$, $z_o$, $z_{oh}$.

### E.3.2 Above the surface layer

Above the surface layer the turbulent fluxes are parametrized following the K-theory and the exchange coefficient $K_\psi$ is given by

$$K_\psi = g\, l_m\, l_\psi \left|\frac{\partial V}{\partial z}\right| f(R_i')$$

where the mixing length $l_\psi$ is $l_h$ for energy and $l_m$ for momentum and $R_i'$ is a modified Richardson number.

The original Richardson is modified in order to include shallow convection effect (which is dependent on the specific humidity) according to Geleyn (1986).

Furthermore, in stable cases, a limitation on the Richardson number has been introduced through the definition of a critical value ($R_{icr}$). This modified Richardson number is computed in the following way:

$$R_i' = \frac{R_i}{(1 + \alpha R_i / R_{icr})^{1/\alpha}} \qquad \text{if } R_i > 0, \text{ stable case}$$

(and otherwise $R_i' = R_i$), where $R_{icr}$ depends on the height according to

$$1/R_{icr}(z) = USURIC\left(1 + (USURICL - 1)\left(\frac{l_h(z)}{k(z + z_{oh})}\right)^{USURICE}\right)$$

Here USURIC, USURICL and USURICE are tunable parameters.

For momentum $\alpha = 1$ and for energy fluxes $\alpha = \frac{3R_i + R_{id}}{R_i + R_{id}}$, where $R_{id}$ depends on the height according to

$$1/R_{id}(z) = USURID\left(1 + \left(\frac{H_m H_h}{(z + z_o)(z + z_{oh})}\right)^2\right)^{USURIDE}$$

where $H_m$ (respectively $H_h$) is the scale height for which $l_m$ (respectively $l_h$) decreases to the asymptotic mixing length for momentum (respectively for energy). USURID and USURIDE are tunable parameters.

The introduction of $R_{icr}$ and $R_{id}$ in the computation of the Richardson number was implemented in order to have realistic simulation of the storm activity, while avoiding the negative feedback between stability and surface cooling.

The functions $f(R_i')$ for momentum and energy are given by

$$f_M(R_i') = \begin{cases} \left(\frac{1}{1 + 2bR_i' / \sqrt{1 + dR_i'}}\right) & \text{if } R_i' > 0, \text{ stable case} \\ 1 - \frac{2bR_i'}{1 + 3bc\left(\frac{l_m}{z + z_o}\right)^2 \sqrt{\frac{(-R_i')}{27}}} & \text{if } R_i' < 0, \text{ unstable case} \end{cases}$$

$$f_H(R_i') = \begin{cases} \left(\frac{1}{1 + 3bR_i' \sqrt{1 + dR_i'}}\right) & \text{if } R_i' > 0, \text{ stable case} \\ 1 - \frac{3bR_i}{1 + 3bc\left(\frac{l_h}{z + z_o}\right)^2 \sqrt{\frac{(-R_i')}{27}}} & \text{if } R_i' < 0, \text{ unstable case} \end{cases}$$

More details about the parametrization of turbulent fluxes used in Arpège are described in Louis (1979), Louis *et al.* (1981) and Geleyn *et al.* (1995).

Table E.2: Values of tunable parameters used in the parametrization of vertical turbulent fluxes

| Parameters | OBS T299 | PHYS T449 |
|---|---|---|
| USURIC | 1.0 | 0.175 |
| USURICL | 4.0 | 1.0 |
| USURICE | 0.5 | 0.5 |
| USURID | 0.042 | 0.1 |
| USURIDE | 1.0 | 0.25 |

## E.4 Cloudiness scheme

In the Arpège model the stratiform (or large scale) condensed water content is a function of the excess of water vapor ($\Delta q_{ex}$):

$$l_s = \begin{cases} l_{smax}(1 - e^{-\alpha_s \Delta q_{ex}/l_{smax}}) & \text{if } \frac{q}{q_{sat}} > r_c) \\ 0 & \text{if } \frac{q}{q_{sat}} \leq r_c \end{cases}$$

and

$$\Delta q_{ex} = q - r_c q_{sat}(T, p) \tag{E.3}$$

where $l_s$ is the specific humidity of the stratiform condensate water, $l_{smax}$ is its maximum value, $q$ and $q_{sat}$ are respectively the specific humidity and the saturation specific humidity, and $\alpha_s$ is a tunable parameter.

Here $r_c$ is the minimum relative humidity necessary to create a cloud, $r_c = f(HUCOE, z)$ in such a way that $r_c$ is 1 at the top and at the bottom of atmosphere, and minimum in middle troposphere. HUCOE is a tunable parameter, which controls the triggering of stratiform clouds. The bigger HUCOE is, the larger both the spatial extension and the amount of condensate of these clouds are.

In order to include a shallow convection effect on cloudiness, the equation $(E.3)$ is replaced by the following one:

$$\Delta q_{ex}^n = \frac{\int_n^{n_b} (q^k - r_c^n q_{sat}^n) W(r^k) dp^k}{\int_n^{n_b} W(r^k) dp^k} \tag{E.4}$$

where $n$ and $k$ are indices of vertical levels, and with $r^k = min\left(\frac{s^k - s^n}{QSSC}, 1\right)$, and $W(r^k) = 1 - 3(r^k)^2 + 2(r^k)^3$.

The vertical integral in (E.4) is performed downwards from the current level $n$ to the level $n_b$ where the difference in the dry static energy reaches the threshold value of $QSSC$, which is a tunable parameter. When $QSSC = 0$, $n_b = n$ and the equation (E.4) becomes (E.3). Setting $QSSC > 0$ permits the clouds to form at the top of PBL.

The convective condensate amount is derived from the convective precipitation:

$$l_c = \begin{cases} l_{cmax}(1 - e^{-\alpha_c g \frac{\partial P_c}{\partial p} \Delta t_{phy}/l_{cmax}}) & \text{if } \frac{\partial P_c}{\partial p} \geq 0 \\ 0 & \text{if } \frac{\partial P_c}{\partial p} < 0 \end{cases}$$

where $l_c$ is the specific humidity of convective condensate water, $P_c$ is the convective precipitation flux, $\Delta t_{phy}$ is the time step and $\alpha_c = QSSUSC$ is a tunable parameter, which controls the amount of condensate water in the convective clouds.

The total condensate water contents $l$ is

$$l = l_s + l_c$$

Finally, the total cloudiness is a function of $l_c$, $l_s$ and $q_{sat}$.

Table E.3: Values of the tunable parameters changed in the cloudiness scheme.

| Parameters | EXP T299 | EXP Phys |
|---|---|---|
| HUCOE | 1.7 | 0.3 |
| QSSC | 1600 | 200 |
| QSSUSC | 1 | 2 |

## E.5 The moisture convergence used in the deep convection parametrization

In the Arpège model the deep convection parametrization is based on the scheme described by Bougeault (1985). In this scheme the deep convection occurs when two conditions are fulfilled:

- there is convergence of humidity at low levels
- the vertical temperature profile is unstable

Moreover, this scheme uses a Kuo-type closure, i.e., it assumes that the available humidity is the sum of the large-scale moisture convergence and the humidity tendency produced by the vertical diffusion, and this available humidity must be equal to the sum of convective precipitation and the detrainment by the environment.

Thus, mathematically the available humidity for deep convection is given by

$$LSCON_f = -\Re\left(V \cdot \nabla q + \omega \frac{\partial q}{\partial p}\right) - g\frac{\partial F_q}{\partial p} \tag{E.5}$$

where $F_q$ is the turbulent flux of $q$ and $\Re$ is modulation factor, which takes into account the mesh size effect.

$$\Re = \frac{1}{(1 + m_p \frac{\Delta x_{ref}}{\Delta x_{eq}})^\gamma}$$

Here $m_p$ is the local map factor, $\gamma$ is a tunable parameter and $\Delta x_{eq} = \frac{2\pi r_a}{3N_s}$, where $r_a$ is the earth radius and $N_s$ is the model truncation. $\Delta x_{ref} = REFLKUO$ when a Kuo-type closure is used in the deep convection parametrization.

However, in Arpège, the available moisture can be computed in two different ways. If $LSRCON = false$, the equation E.5 is used. But, if $LSRCON = true$ then the available moisture to deep convection is given by

$$LSCON = LSCON_f - g\frac{\partial P_{LS}}{\partial p}$$

where $P_{LS}$ is the large scale precipitation flux.

The tuning parameter GCVPSI determines the way turbulent fluxes (enthalpy and humidity) are compensated in the convection scheme. It allows a continuous transition between the non-averaging state (GCVPSI=1) to the complete averaging one (GCVPSI=0). A completely local compensation (and not any more a partially "smoothed" one) of turbulent fluxes by convective fluxes has been chosen to reduce the cyclogenetic tendencies (frequency and intensity) at various scales.

Table E.4: Values of the parameters changed in the closure of the deep convection scheme.

| Parameters | EXP T299 | EXP Phys |
|---|---|---|
| REFLKUO | 10000 | 5000 |
| LSRCON | TRUE | FALSE |
| GCVPSI | 0.5 | 1 |

# Appendix F

## The use of an ensemble approach to study the Background Error Covariances in a Global NWP model

***(paper to appear in Monthly Weather Review)***

M. Belo Pereira*[1] and L. Berre[2]
*Instituto de Meteorologia - DVM/VPN, Lisboa, Portugal
[2]Météo-France - CNRM/GMAP, Toulouse, France

ABSTRACT

The estimation of the background error statistics is a key issue for data assimilation. Their time average is here estimated by using an analysis ensemble method. The experiments are performed with the nonstretched version of the Arpège global model, in a perfect-model context.

The global (spatially averaged) correlation functions are sharper in the ensemble method than in the so-called NMC method. This is shown to be closely related to the differences in the analysis step representation.

The local (spatially varying) variances appear to reflect some effects of the data density and of the atmospheric variability. The resulting geographical contrasts are found to be partly different from those that are visible in the operational variances and in the NMC method.

An economical estimate is also introduced to calculate and compare the local correlation length scales. This allows to diagnose some existing heterogeneities and anisotropies. This information can be useful also for the modeling of heterogeneous covariances, based e.g. on wavelets.

The implementation of the global covariances and of the local variances, which are provided by the ensemble method, appears moreover to have a positive impact on the forecast quality.

[1]email: margarida.belo@meteo.pt

## F.1 Introduction

Most of the major operational NWP centers use variational methods to combine observations with a background state provided by a short-range forecast, in order to produce an optimal estimate of the atmospheric state. However, this estimate is only optimal if both the observation and background error covariance matrices are correctly specified in the analysis. Therefore, an accurate specification of these covariance matrices is very important for the quality of the assimilation system. However, the estimation of the background error statistics is not straightforward, since the truth is never exactly known. In an NWP operational context, the so-called NMC method is an example of one possible approach used to estimate the background error covariances. This method was first implemented in the National Meteorological Center (nowadays named National Center for Environmental Prediction) (Parrish and Derber 1992). This approach is based on an empirical assumption: the differences between forecasts of different lengths, but valid at the same time, have similar structures to those of the short range forecast errors. This method has the advantages of being easy to implement and of providing an estimate for the whole globe. This explains probably why this method was or is operational in many NWP centers: at ECMWF (Rabier *et al.* 1998), at CMC in Canada (Gauthier *et al.* 1999), at Météo-France (Desroziers *et al.* 1995), at the UKMO (e.g. Lorenc *et al.* 2000). This method was used in limited area models too (Berre 2000; Gustafsson *et al.* 2001).

More recently, an interesting alternative technique was proposed by Houtekamer *et al.* (1996). It is based on an ensemble of assimilation experiments: for each ensemble member, the observations are randomly perturbed, as well as the physical parametrizations. This method has been also implemented operationally at ECMWF (Fisher, 2003). This ensemble approach has also been tested and studied at Météo France in the Arpège global model and in the Aladin limited area model (Ştefănescu *et al.* 2006, Berre *et al.* 2006). The present paper describes the results that were obtained for the Arpège global NWP model. A perfect-model framework has been used for the ensemble experiments in order to simplify the interpretation of the first results. This will also allow the investigation of the role of the model error at a later stage.

The resulting covariance estimates have been compared with those that were previously operational, by performing both diagnostic and impact studies: the global (i.e. spatially averaged) ensemble covariances have been compared with those of the NMC method, and the local (i.e. spatially varying) ensemble variances have also been compared with the local operational variances (the latter correspond to a static 3D map of variances, that reflects some typical data density contrasts). Some emphasis has been given to study the role of the analysis step in this context.

Some significant effort has been devoted to the diagnosis and examination of the local ensemble correlations. This has been done by developing an economical method to estimate the local correlation length scales. The study of these local correlations gives interesting informations about the existing heterogeneities and anisotropies. This can be also useful for future evaluation of heterogeneous covariance formulations, such as those based on wavelets (Fisher (2003), Deckmyn and Berre (2005)) or recursive filters (Wu *et al.* 2002).

The structure of the paper is the following. The formalism of the ensemble simulation of the model state error is presented in section F.2, where it is compared with the implicit formalism of the NMC method. Section F.3 concerns the comparison between the global covariance estimates of ensemble and NMC methods. The local covariance estimates are diagnosed in section F.4. The impact studies of the ensemble global covariances and of the ensemble local variances are then described in section F.5.

## F.2 The formalism of the ensemble simulation of the error evolution

In this section, the formalism of the method based on an ensemble of assimilation experiments is briefly described. It will be also compared with the implicit formalism of the NMC method. A linear framework is considered to make the derivations simpler. The reader is referred to Berre *et al.* (2006) for more detailed derivations (a formalism in a non linear framework is described in Žagar *et al.* (2004)). In the current paper we summarize the main features that are important for the interpretation of the experimental results. This concerns for instance the differences between the two methods regarding the analysis step representation.

The exact evolution of the model state error is first introduced. The way this error evolution is simulated in the ensemble approach is then explained. It is finally contrasted with the corresponding implicit formalism of the NMC method. The true errors will be noted $e$, while the simulated errors will be noted $\varepsilon$ and $\epsilon_{NMC}$ respectively in the ensemble and NMC methods. The experimental environment of the ensemble runs is finally described in the last part of the current section.

### F.2.1 The exact model state error evolution

The involved assimilation experiments correspond to a succession of analyses and forecasts. This process implies a specific evolution for the model state errors.

Starting from an uncertain analysis $x_a^i$ (valid at time $t_i$), one will obtain an uncertain six-hour (6h) forecast $x_b^{i+1}$ (valid at time $t_{i+1} = t_i + 6h$) by integrating the NWP model. The evolution of the analysis errors $e_a^i$, into the background errors $e_b^{i+1}$, corresponds to the following equation:

$$e_b^{i+1} = \mathbf{M} e_a^i + e_m^{i+1}$$

where $\mathbf{M}$ is the six-hour forecast operator, and $e_m^{i+1}$ is the accumulated model error (during the six-hour period).

The six-hour forecast field $x_b^{i+1}$ will then be used as the background for a new analysis (that is valid at time $t_{i+1}$, and which will use the observation vector $y^{i+1}$). The evolution of the background errors $e_b^{i+1}$ into the analysis errors $e_a^{i+1}$ corresponds to the following equation:

$$e_a^{i+1} = e_b^{i+1} + \mathbf{K}(e_o^{i+1} - \mathbf{H} e_b^{i+1}) = \mathbf{K} e_o^{i+1} + (I - \mathbf{KH}) e_b^{i+1} \qquad \text{(F.1)}$$

where $\mathbf{K}$ is the classical gain matrix ($\mathbf{K} = \mathbf{BH}^T(\mathbf{HBH}^T + \mathbf{R})^{-1}$, where $\mathbf{B}, \mathbf{R}$ are the specified background and observation error covariance matrices), $e_o^{i+1}$ is the observation error (which is itself the sum of the measurement error and of the representativeness error), and $\mathbf{H}$ is the observation operator.

### F.2.2 The ensemble simulation of the error evolution

The method based on an ensemble of assimilation experiments may be seen as a technique that intends to simulate the exact error evolution. This can be illustrated by deriving the equations of evolution of the ensemble dispersion (which aims to simulate the involved uncertainties) during a basic forecast/analysis step. The ensemble method is illustrated schematically in figure F.1.

Considering that at a given time an ensemble of (perturbed) analyses is available, from this ensemble one can obtain an ensemble of (perturbed) six-hour forecasts, by integrating the NWP model. The analysis differences ($\varepsilon_a^i$) between different perturbed analyses of the ensemble simulate the analysis error ($e_a^i$). The evolution of $\varepsilon_a^i$ into the background differences $\varepsilon_b^{i+1}$ corresponds to the following equation:

$$\varepsilon_b^{i+1} = \mathbf{M} \varepsilon_a^i + \varepsilon_m^{i+1}$$

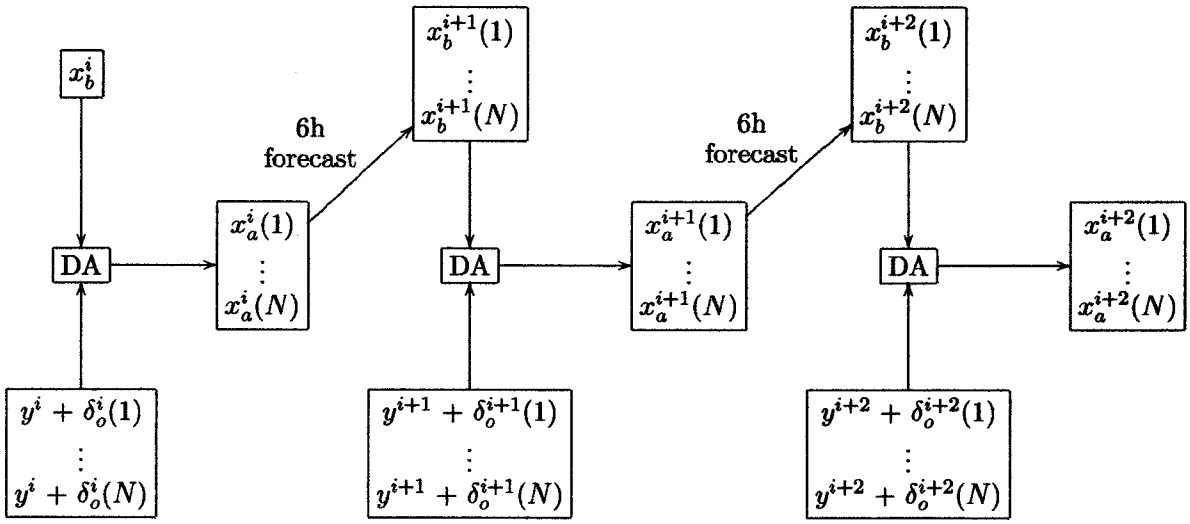


Figure F.1: Schematic illustration of the ensemble method, where DA stands for data assimilation. For each member $n$ (for $n = 1, ..., N$), $\delta_o^i(n)$ represents the random realization of the observation perturbations at time $t_i$, and the background field valid at $t_{i+1} = t_i + 6h$ is $x_b^{i+1}(n) = \mathbf{M}x_a^i(n)$. The analysis for each member is $x_a^{i+1}(n) = x_b^{i+1}(n) + \mathbf{K}(y^{i+1}(n) - \mathbf{H}x_b^{i+1}(n))$. For the first analysis step, the operational (unperturbed) background is used, for all ensemble members.

where $\varepsilon_m^{i+1}$ represents the simulated model error, which results from some model perturbations that can be added.

Similarly, the background differences $\varepsilon_b^{i+1}$ will be evolved into analysis differences $\varepsilon_a^{i+1}$, according to the following equation:

$$\varepsilon_a^{i+1} = \varepsilon_b^{i+1} + \mathbf{K}(\varepsilon_o^{i+1} - \mathbf{H}\varepsilon_b^{i+1}) = \mathbf{K}\varepsilon_o^{i+1} + (I - \mathbf{KH})\varepsilon_b^{i+1} \tag{F.2}$$

where $\varepsilon_o^{i+1}$ are observation differences (between the ensemble members) that simulate the observation errors $e_o^{i+1}$. For each ensemble member, the observations are perturbed by adding independent random Gaussian perturbations $\delta_o$, which are drawn from the specified observation error covariance matrix $\mathbf{R}$.

It is important to notice that the basic form of the last two equations and the involved operators are the same as in the exact error evolution. Thus, the ensemble approach has the advantage of simulating two basic and important components of the exact error evolution: the forecast evolution provided by the model (represented by $\mathbf{M}$), and the analysis step (represented by $\mathbf{K}$ and $\mathbf{I} - \mathbf{KH}$, which are the respective analysis weights of the observations and of the background). In practice, these two components are simulated by applying a succession of analyses and forecasts to some perturbed states.

Finally, one may mention that with the ensemble approach the model errors $e_m$ can be simulated, either by using different NWP models for the different ensemble members (Houtekamer *et al.* 1996), or by drawing random realizations from an estimate of the model error covariance matrix (Mitchell *et al.* 2002).

In the case of a perfect model framework, $\varepsilon_m$ is simply set to zero. This approach is expected to lead e.g. to an underestimation of the error variances. Nevertheless, it is a simpler approach when starting the experimentations. It also gives the possibility to assess the impact of the possible future model error simulations.

### F.2.3 The NMC simulation of the error evolution

The NMC method may also be seen as a technique that simulates the evolution of the model state errors. One may consider the case where differences are taken between 36h and 12h

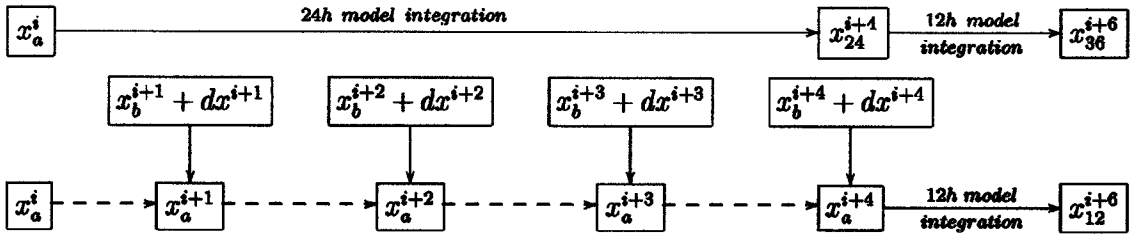


Figure F.2: Schematic illustration of the NMC method. $x_{24}^{i+4}$ is the 24h forecast that is valid at time $t_{i+4}$. Similarly, $x_{36}^{i+6}, x_{12}^{i+6}$ are respectively the 36h and 12h forecasts that are valid at time $t_{i+6}$. (See the text for the other notations.)

forecasts that are valid at the same time (see figure F.2). It may be noted that there is a 24h period between the respective starting dates of these two forecasts. During this period, there is a succession of analyses and 6h forecasts. After this 24h period, there is a final forecast step (over a 12 h period).

In order to show the main differences between the NMC approach and the exact and ensemble error evolutions, it is useful to write the equations for the first and for the final analysis steps in the NMC method.

For the first analysis step of the NMC method, the simulated analysis error $\epsilon_{a,NMC}^{i+1}$ is simply equal to the analysis increment:

$$\epsilon_{a,NMC}^{i+1} = \mathbf{K}e_o^{i+1} - \mathbf{KH}e_b^{i+1}$$

Comparing this equation with equation (F.1), it is possible to summarize important features of the NMC method.

Regarding the estimated variances of the analysis errors, in a similar way as discussed in Bouttier (1994), one may consider the following three limit cases:

- In regions where the data density is high, and the observation errors have similar variances and similar spatial correlations as the background errors, **KH** and **K** are close to $I/2$. This implies that $e_a^{i+1} \approx \frac{1}{2}(e_o^{i+1} + e_b^{i+1})$, while $\epsilon_{a,NMC}^{i+1} \approx \frac{1}{2}(e_o^{i+1} - e_b^{i+1})$. The consequence is that, in this case, the variances of the analysis increments are similar to the variances of the exact analysis errors. (Note that the variances of $e_o$ and $e_b$ are added, rather than subtracted: $\langle(\epsilon_{a,NMC}^{i+1})^2\rangle \approx \frac{1}{4}(\langle(e_o^{i+1})^2\rangle + \langle(e_b^{i+1})^2\rangle)$.

- In regions where the data density is high and the observations have a high quality, **KH** and **K** are close to $I$. This implies that $e_a^{i+1} \approx e_o^{i+1}$, while $\epsilon_{a,NMC}^{i+1} \approx e_o^{i+1} - e_b^{i+1}$. In other words, the variances of the analysis increments are likely to be an overestimation of the variances of the exact analysis errors.

- In regions where the data density is low and the (few) observations have a low quality, **KH** and **K** are close to 0. This implies that $e_a^{i+1} \approx e_b^{i+1}$, while $\epsilon_{a,NMC}^{i+1} \approx 0$. In other words, in data sparse areas, the simulated analysis error variances are likely to be underestimated by the NMC approach.

Regarding the estimated correlations of the analysis errors, one may note that the exact operator $(I - \mathbf{KH})$ is replaced by **KH**. If the observation error covariances are similar to the background error covariances (as evoked in the first limit case above), these two operators are similar, and the analysis error correlations can be expected to be well approximated. However, the observation errors are usually less spatially correlated than the background errors: this

implies that $(I - \mathbf{KH})$ acts as a high-pass filter, while $\mathbf{KH}$ is rather a low-pass filter (e.g. Daley (1991), section 4.5). The consequence is that in this case the analysis error correlations are likely to be overestimated by the NMC method.

For the "final" analysis step (which corresponds to the fourth analysis step in the 24 h period) in the NMC method, the simulated analysis error is:

$$\epsilon_{a,NMC}^{i+4} = \epsilon_{b,NMC}^{i+4} + dx^{i+4} = (\sum_{k=1}^{3} \mathbf{M}^{4-k}\, dx^{i+k}) + dx^{i+4} \tag{F.3}$$

where $\epsilon_{b,NMC}^{i+4}$ can be seen as the NMC background perturbation, which contains the contributions of the earlier analysis increments, that were evolved in time and accumulated.

Thus, in the NMC method, the analysis step consists in adding the (unperturbed) analysis increment to a background perturbation (which corresponds itself to some earlier evolved increments). In contrast, in the exact and ensemble evolutions, the analysis step consists in applying the analysis equation to a vector of background differences and to a vector of observation differences (the differences are the "true errors" in the exact evolution).

The last equation can be re-written as:

$$\epsilon_{a,NMC}^{i+4} = \sum_{k=1}^{4} \mathbf{M}^{4-k}\, dx^{i+k}$$

In addition, in the NMC method, this final analysis perturbation is evolved during 12 additional hours, which provides the final 36h-12h forecast differences, $\epsilon_{b,NMC}^{i+6}$ (that intends to simulate $e_b^{i+6}$):

$$\epsilon_{b,NMC}^{i+6} = \mathbf{M}^2\, \epsilon_{a,NMC}^{i+4} = \mathbf{M}^2\, (\sum_{k=1}^{4} \mathbf{M}^{4-k}\, dx^{i+k})$$

which can be compared with the corresponding equation for the background differences in the ensemble method:

$$\varepsilon_b^{i+6} = \mathbf{M}\, \varepsilon_a^{i+5}$$

The comparison between the last two equations allows to summarize the main three peculiarities of the NMC method, compared with the ensemble method:

- the involvement of longer forecast ranges (see the occurrence of the matrices $\mathbf{M}^2, \mathbf{M}^{4-i}$ instead of the 6h matrix $\mathbf{M}$) ;
- the accumulation of several increments (see the occurrence of the operator $\sum$) ;
- the involvement of analysis increments $dx$, instead of analysis differences $\varepsilon_a$.

The involvement of longer forecast ranges has been evoked by Fisher (2003), to explain the differences between the NMC and ensemble methods. We will show in the section F.3.3 that the differences in the representation of the analysis step (i.e. the third item in the list above) also play an important role.

Although the previous derivations are presented in a linear framework, it should be also mentioned that, in practice, both the ensemble and NMC methods involve a non-linear model during the forecast step, and a non-linear analysis scheme (such as 4D-Var) may be used for the analysis step. This allows to represent some non-linear effects in the simulation of the error evolution, such as the error growth saturation in the small scales during the forecast step.

### F.2.4 The experimental framework and the operational covariances

The results shown in the present paper were produced from one ensemble with five 4D-Var assimilation cycles, that were performed from 1 February to 24 March 2002. The first analysis ensemble is created by randomly perturbing the observations, and by using the (unperturbed) operational background field. The model integration will then provide an ensemble of backgrounds (see figure F.1). Thus, after the first analysis step, the perturbed analyses result from both observation and background perturbations. The amplitude of the background differences grows from zero to stable values during the first three days. This preliminary period is therefore not included in the statistics, which are rather calculated from 04/02 to 24/03 (which is a 49 day period).

The members were arbitrarily numbered from 1 to 5 $(x_b(1), x_b(2), x_b(3), x_b(4), x_b(5))$, from which four sets of member differences were calculated $(x_b(1) - x_b(2), x_b(2) - x_b(3), x_b(3) - x_b(4), x_b(4) - x_b(5))$. So, the correlations and standard deviations of the background error are computed from $4 \times 49 = 196$ differences between 6h forecasts, for each of the four daily analysis times (00, 06, 12 and 18 UTC). Mostly, the results from fields valid at 18UTC will be shown.

It may be mentioned that using differences between background states from separate perturbed experiments is equivalent to the use of differences between a perturbed background and the background from an unperturbed control experiment, except for a factor 2 in the implied covariances (see e.g. Berre *et al.* (2006): it can be shown that the covariance of the difference between two perturbed backgrounds is equal to twice the covariance of the single background perturbations).

The experiments have been performed with the Arpège global model (in its non-stretched version, i.e. with a uniform resolution), and its 4D-Var scheme (Rabier *et al.* (2000); Veersé and Thépaut (1998)). The model has been integrated at T299 triangular truncation with 41 levels. The formulation of **B** matrix that is used in the Arpège 4D-Var is described in Derber and Bouttier (1999). The global covariances were those calculated by using the NMC method. In this formulation, vorticity plays an important role, as its statistics determine the main part of the wind covariances, and also the main part of the temperature and surface pressure covariances. The vorticity correlations do not vary horizontally, but the vorticity standard deviations are allowed to vary geographically.

Previous studies (Rabier *et al.* (1998), Bouttier (1994) and McNally (2000)) indicate that the geographical variations of the NMC standard deviations are not expected to be realistic. Therefore, at ECMWF, a cycling algorithm is used to determine space and time variations of these vorticity standard deviations, as described in the appendix of Derber and Bouttier (1999). This cycling algorithm is not implemented in the operational Arpège 4D-Var (at Météo France). Instead, the ECMWF maps that were produced for 20 October 1998 are used. As will be illustrated in section F.4.1, these maps allow to represent e.g. some typical data density contrasts over the globe.

Since in the operational Arpège 4D-Var the average vertical profile of the standard deviation is provided by the NMC method, only the normalized values of vorticity standard deviations given by the cycling technique are used. In other words, the cycling algorithm is used only to describe the horizontal deviations from the NMC average standard deviation (at each vertical level).

Some randomization techniques (Andersson and Fisher 1998) allow moreover to deduce some maps of standard deviations for wind, temperature and geopotential. These latter maps are used in the quality control of the Arpège 4D-Var, when a first-guess check is performed to reject observations whose values are too different from the background values.

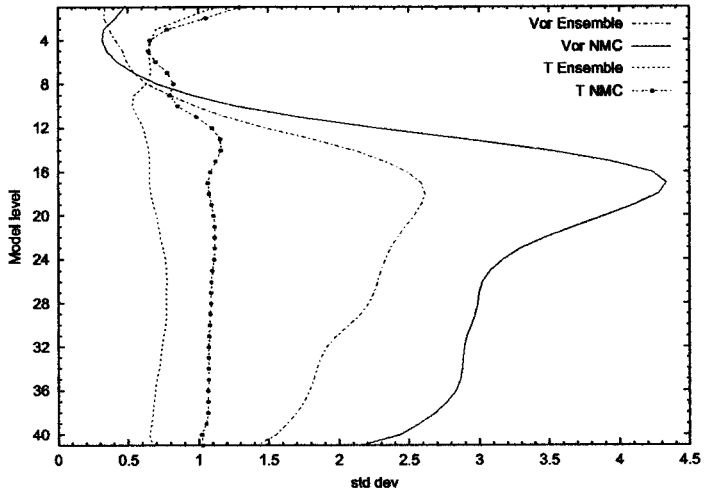


Figure F.3: Vertical profile of standard deviation of the vorticity ($10^{-5}s^{-1}$) and temperature (K) background errors.

## F.3 Global background error covariances

In this section we present the comparison between the ensemble and NMC methods, relatively to the estimation of the global background error covariances.

The covariances were computed in spectral space at truncation T179 and the results will be presented for the following variables: vorticity, temperature and logarithm of surface pressure ($P_s$) (referred to "surface pressure" in the remainder of the paper).

The influence of the analysis step on these differences will be studied in the final part of this section.

### F.3.1 Standard deviations

The standard deviations of temperature and vorticity background errors estimated by the NMC and ensemble methods is presented in figure F.3. One can notice that there are some similarities between the two methods. For instance, both methods indicate that the background errors of temperature show small variations with height in troposphere. Another common feature is that the background errors of vorticity increase significantly with height, with a maximum near the tropopause. This maximum is related to the jet winds in the middle latitudes. Nevertheless, the contrast between the jet level and the low levels is less strong in the ensemble method than in the NMC method.

The standard deviations are obviously larger in the NMC method than in the ensemble method. The absence of model error simulation in the current version of the ensemble experiments is likely to contribute to this difference. Moreover, the accumulation of four analysis increments and the involvement of long forecast ranges in the NMC method contribute probably also to this result.

### F.3.2 Horizontal correlations

The correlation spectra will be shown instead of the covariance spectra, since this allows to compare more directly the relative contributions of the different horizontal scales, in the two methods. The figure F.4 presents the correlation spectra for temperature. One can see that compared with the NMC method, the ensemble approach emphasizes the relative contributions of the small scales (i.e. compared with the contributions of the large scales). This result is

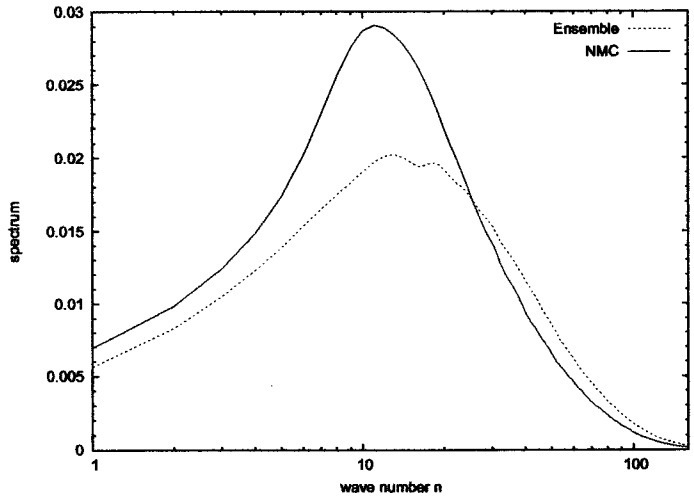


Figure F.4: Autocorrelation spectra of the background error of temperature at level 21 (near 500hPa) for NMC (full line) and ensemble (dotted line) methods.

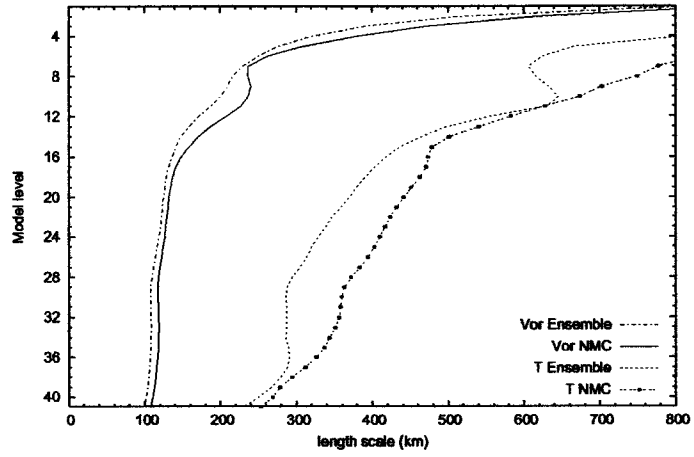


Figure F.5: Horizontal length scale of the autocorrelation function of the vorticity and temperature background error, estimated by the NMC and ensemble method.

also found for the other variables. As illustrated in figure F.5, this implies that the corresponding correlation length scales (as defined by Daley (1991, p.110) and Rabier *et al.* (1998)) are smaller in ensemble than in NMC method. The usual increase of length scale with height can also be identified in figure F.5, in accordance with the increase with height of the large scale contributions.

The enhancement of the small scale horizontal contributions implies also that the vertical correlations are sharper in the ensemble method (see section F.4.3). Another consequence is that the contributions of the unbalanced components are emphasized in the multivariate formulation (not shown). All these results are consistent with those mentioned by Fisher (2003).

### F.3.3 The influence of the analysis step

In section F.2, it has been noticed that the ensemble and NMC methods differ regarding the representation of the analysis step. The ensemble method simulates the reduction of the model state errors when combining an uncertain background with uncertain observations. To contrast with this, the NMC method is rather relying on the accumulation and time evolution of four

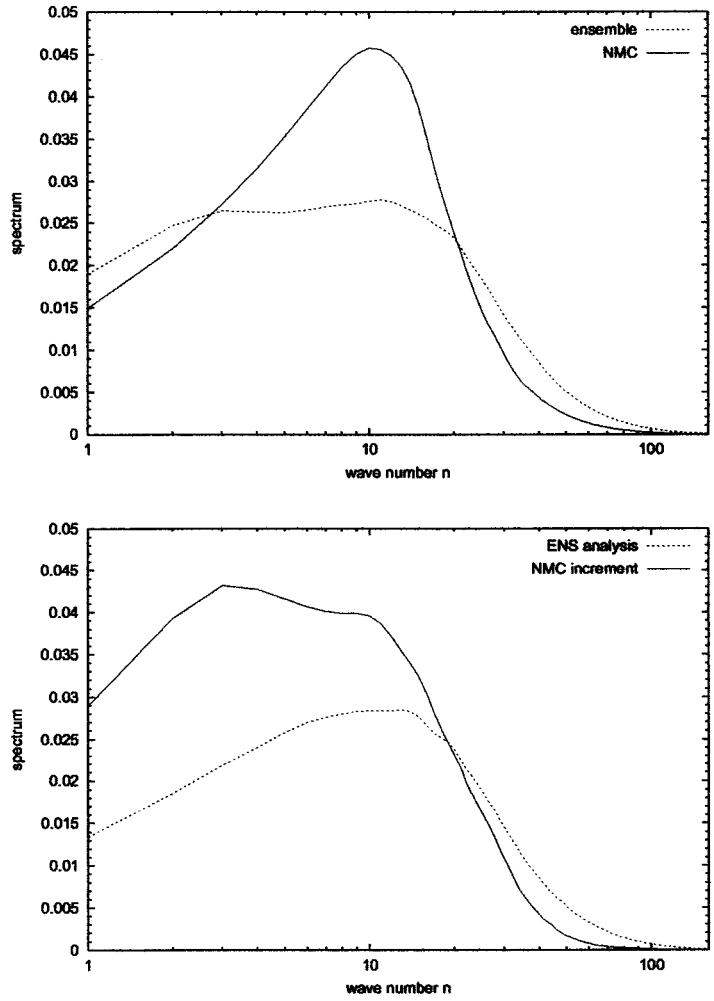


Figure F.6: Autocorrelation spectra for surface pressure. Top panel: Background error spectra estimated by the ensemble (dashed line) and the NMC method (solid line). Bottom panel: spectra of the analysis dispersion in the ensemble method (dashed line) and of the analysis increment in the NMC method (solid line).

successive analysis increments.

An indication of the influence of these analysis representation differences is shown in figure F.6. The top panel represents the background error correlation spectra of $P_s$, estimated by the ensemble and NMC methods. The contributions of the small scales and of the planetary scales are emphasized in the ensemble method.

The bottom panel corresponds to the comparison between the correlation spectra of the analysis increment (in the NMC method) and of the analysis dispersion (in the ensemble method). It appears that the analysis increment spectrum is much larger scale than the analysis dispersion spectrum. Such a difference is consistent with the expectation that the spatial correlations of the analysis increment are an overestimation of the analysis error correlations (see section F.2.3). As the analysis increment is one fundamental ingredient of the NMC method, this contrast explains the larger relative contributions of the small scales to the background errors in the ensemble method. This indicates that the analysis representation differences play an important role in the scale differences between the two methods.

## F.4 Local background error covariances

In this section, the geographical variations of the ensemble covariances are studied. The ensemble local standard deviations will be compared with the operational local standard deviations, and also with the NMC local standard deviations.

The geographical variations of the ensemble correlations will also be examined. This will be done in particular by introducing an economical estimation of the local correlation length scales. Concerning this aspect a comparison will be also done with the NMC method.

Finally, the corresponding differences in the vertical correlations and in the analysis/background errors will be also illustrated.

### F.4.1 Standard deviations

All the statistics presented in this subsection (as in the whole section) were computed on a $1.5 \times 1.5$ latitude-longitude grid (which is comparable to a truncation T127 approximately). However, the standard deviations maps presented here have been truncated at T21 for vorticity and at T79 for temperature. These truncations have been chosen in order to filter the sampling noise, while retaining the relevant features.

The figure F.7 shows the maps of normalized standard deviations of vorticity background error at 500hPa, estimated by the ensemble method and used in the operational experiments. (The normalization consists simply in a division by the horizontal average of the standard deviation: the values in the figures correspond to some horizontal modulations, that are to be applied to the global covariances (see section F.2.4)).

Some common features can be identified, e.g. regarding the latitudinal variations and the influence of the data density: the standard deviations are smaller in the tropics than in the middle latitudes, and they are also relatively small over data rich regions, such as Europe and the United States of America (USA).

On the other hand, in the ensemble method, the local minima are more marked over the tropics, and less pronounced over the data rich areas. In addition, the local maxima over the mid-latitude oceans are enhanced, such as over the Northern Pacific, the Northern Atlantic, and the circumpolar ocean of the Southern Hemisphere. The ensemble method provides also some local maxima near the North Pole and over the Himalayas, which are not present in the operational map.

The geographical variations of the ensemble standard deviations have been also compared with those provided by the NMC method. The results for temperature at the levels 37 (near 970hPa) and 27 (near 700hPa) will be used to illustrate some typical differences (figures F.8 and F.9, respectively).

Near 970hPa, some similar latitudinal variations are found according to the two methods. For instance, the mid-latitude values are larger than in the tropics, and some large values around the North Pole are visible in the two estimates.

One of most striking differences is the occurrence of some strong maxima in the central regions of North America, according to the NMC method. When the ensemble method is used, the background errors have a local maximum in Canada, near the Hudson bay, but they have relatively small values in USA. A local minimum over Europe is also more visible in the ensemble method than in the NMC method.

Over e.g. Europe and North America, the differences between the two methods can be partly explained by the tendency of the NMC method to overestimate the background errors in data rich areas, as explained in section F.2.3. On the other hand, it may be that the perfect model assumption leads to an underestimation of the background errors in some regions by the ensemble method. For instance, one could consider that the temperature errors in the Boundary Layer in mountainous areas (e.g. Rocky mountains and Alps) might be underestimated by the ensemble method.

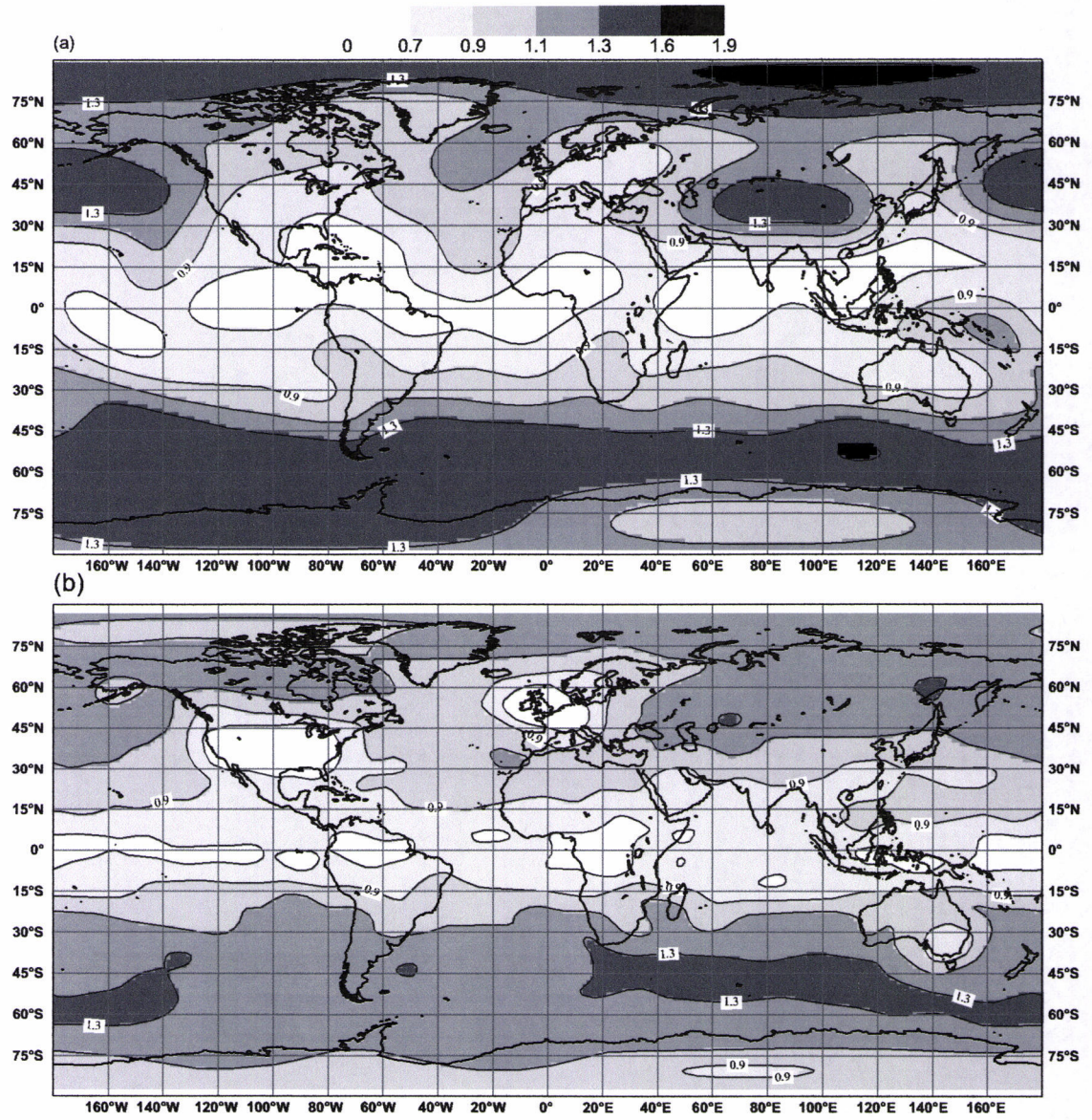


Figure F.7: Normalized standard deviations of the vorticity background error (around 500 hPa), (a) estimated by the ensemble method and (b) used in the operational experiments. (The bottom map has truncation T42, as used in the former operational 4D-Var).

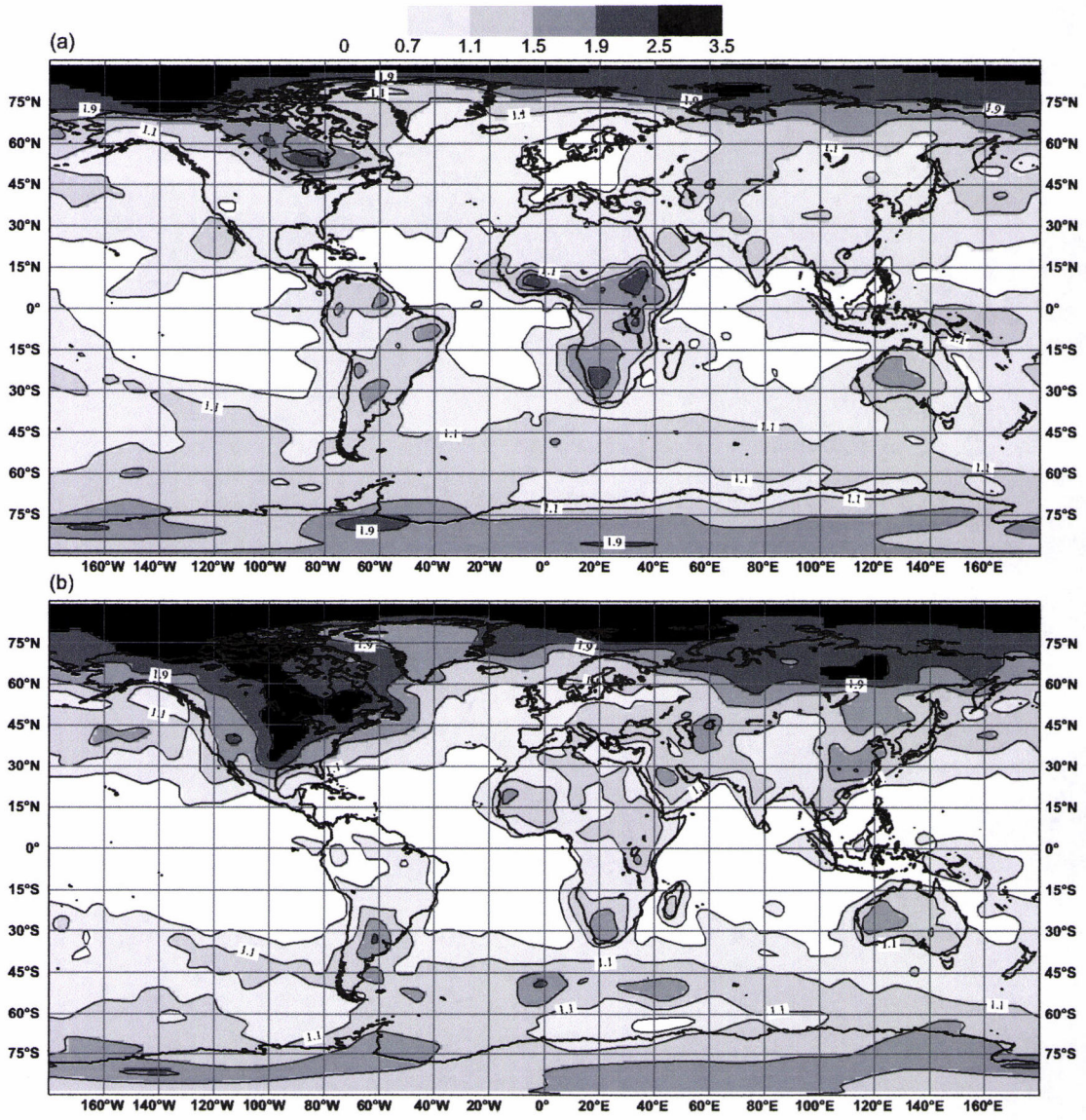


Figure F.8: Normalized standard deviations of the temperature background error at level 37 (i.e. around 970 hPa), using the ensemble method (top) and the NMC method (bottom).

The figure F.8 also indicates that, according to the ensemble method, large values of background errors occur in Equatorial Africa. This is consistent with the large atmospheric variability observed in this area, associated with the Inter-Tropical Convergence Zone (ITCZ). Such a feature is not well captured by the NMC method. More generally, in the tropics, the magnitude of the background error is larger when estimated by the ensemble method than by the NMC method. These differences may be explained by the fact that the NMC method tends to underestimate the background errors in data sparse areas, as mentioned before.

Near 700hPa, when the background errors are estimated by the ensemble method, the contrast between data rich and data sparse regions is very clear (figure F.9). Generally, the NMC method is not able to represent such a feature. Moreover, in the Southern Hemisphere, the largest background errors of the ensemble method are located in the Tropical Eastern Pacific, and in the Tropical Atlantic. This corresponds to areas of marine stratocumulus. This may reflect that some large uncertainties exist in these regions: these uncertainties may be related to some strong sensitivity of the shallow convection and radiation parametrizations, with respect to the initial conditions of the NWP model. The NMC method also reflects such uncertainties,

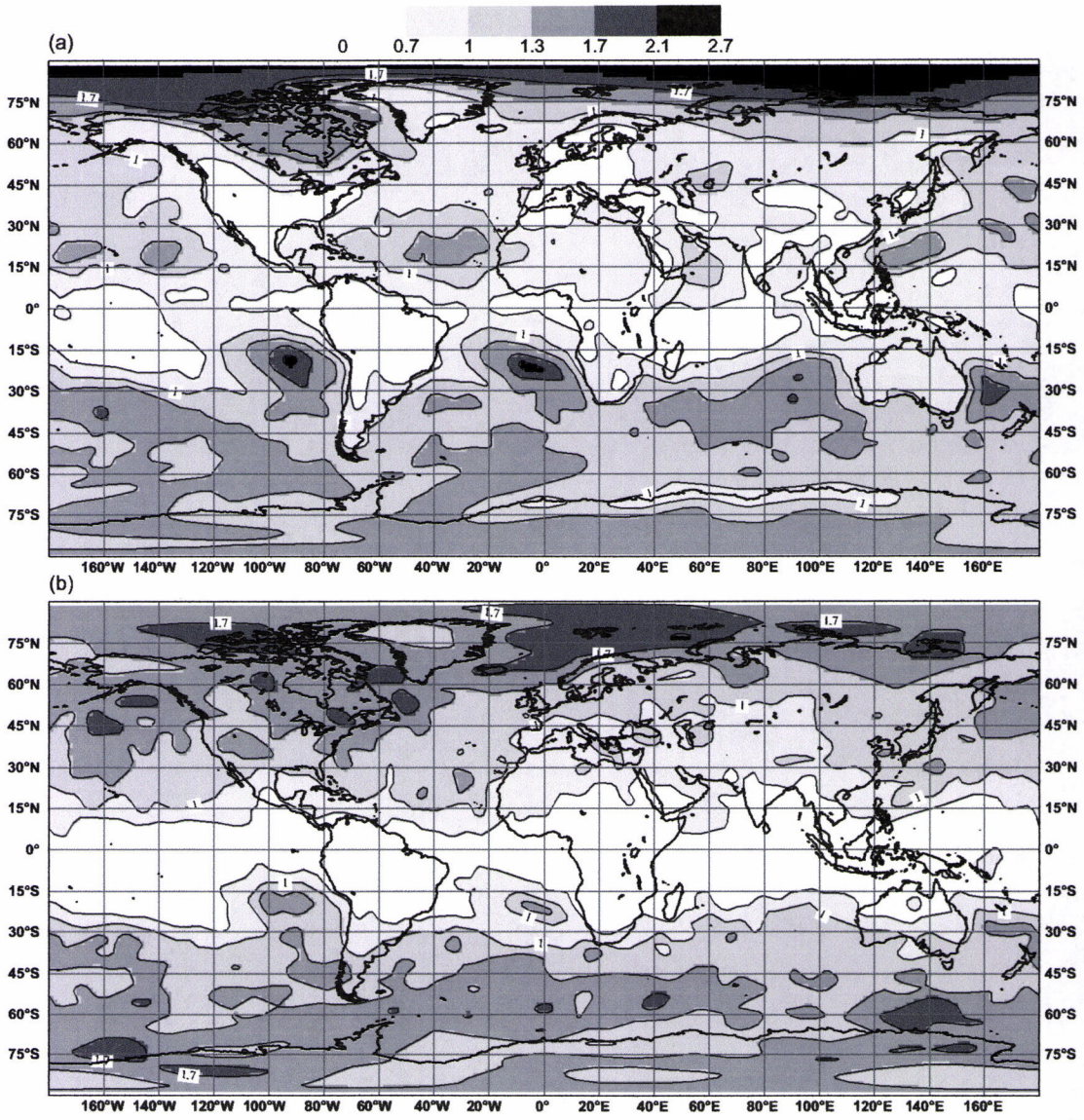


Figure F.9: Normalized standard deviations of the temperature background error at level 27 (i.e. around 700 hPa), using the ensemble method (top) and the NMC method (bottom).

but less clearly than the ensemble method. This may be due to an underestimation of analysis errors (by the NMC method) over these data-poor oceanic regions.

### F.4.2 Horizontal length scales

Diagnosing the full geographical variations of the correlations is a significant challenge: the total number of correlation values to be calculated is the square of the number of gridpoints. This is much larger than what is required for e.g. the calculation and the representation of the geographical maps of standard-deviation. In this context, it is meaningful to introduce an economical estimation of the local correlation length scale. This has been achieved by using the background error variance of each variable and of its derivative.

For any simulated error $\epsilon$, it is indeed possible to define a zonal length scale ($L_x^\epsilon$) and a

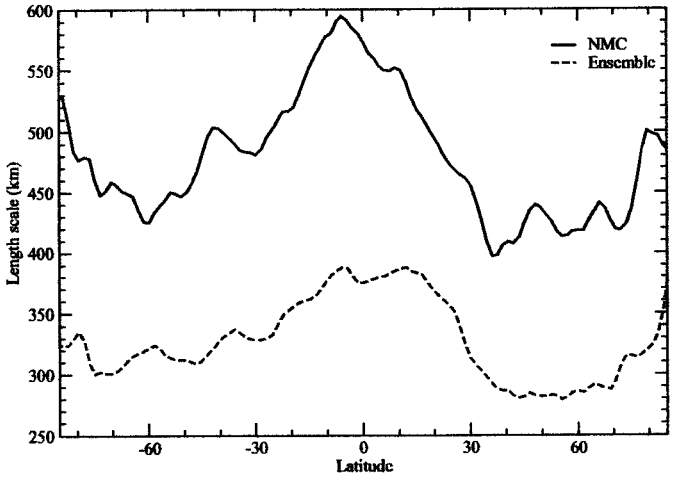


Figure F.10: North-south variation of the length scale of surface pressure, estimated by the ensemble (dashed line) and NMC (full line) methods.

meridional length scale ($L_y^\epsilon$) as (see appendix for derivations):

$$L_x^\epsilon = \sqrt{\frac{\sigma^2(\epsilon)}{\sigma^2(\frac{\partial \epsilon}{\partial x}) - \left(\frac{\partial \sigma(\epsilon)}{\partial x}\right)^2}}, \qquad L_y^\epsilon = \sqrt{\frac{\sigma^2(\epsilon)}{\sigma^2(\frac{\partial \epsilon}{\partial y}) - \left(\frac{\partial \sigma(\epsilon)}{\partial y}\right)^2}} \tag{F.4}$$

where $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ are respectively the derivatives in $x$ and $y$ directions, $\sigma^2(\epsilon)$ and $\sigma^2(\frac{\partial \epsilon}{\partial x})$ are respectively the variances of the background error $\epsilon$ and of the derivative $\frac{\partial \epsilon}{\partial x}$. The 2D correlation length scale is given by $L^\epsilon = \sqrt{({L_x^\epsilon}^2 + {L_y^\epsilon}^2)/2}$.

This method of computing $L$ is economical: for a domain with $N$ gridpoints, one needs to calculate $3N$ variances, instead of computing $N^2$ correlations or covariances. Moreover, it provides a synthetic view of the geographical variations of the correlation functions.

In order to validate this estimate of the length scales, the global length scales have been compared with the ones computed in spectral space. It was found that the length scales estimated in gridpoint space are similar, but slightly larger than those estimated in spectral space, which is consistent with the different resolutions used in the two spaces.

The length scale estimates that are provided by the ensemble and NMC methods will be noted $L_{ENS}^\epsilon$ and $L_{NMC}^\epsilon$ respectively. The latitudinal variations of $L_{NMC}^\epsilon$ and $L_{ENS}^\epsilon$ were examined, for different meteorological variables. According to both methods, the correlation length scale is larger in the tropics than in the middle and high latitudes, for all the variables. This is illustrated in figure F.10 for surface pressure. Following the physical explanation from Lindzen and Fox-Rabinovitz (1989), Ingleby (2001) suggested that the increase of the horizontal correlation length scale, when approaching the equator, reflects the latitudinal dependence of the Rossby radius of deformation.

From figure F.10, it is also clear that $L_{ENS}^\epsilon$ is much shorter than $L_{NMC}^\epsilon$. Moreover, in the middle latitudes, the correlation length scale of $P_s$ is smaller in the Northern Hemisphere than in the Southern Hemisphere. The global map of the local length scales of $P_s$ (not shown) indicates that this is related to the larger presence of land surfaces in the Northern Hemisphere.

The length scales for different regions of the globe were also examined. The corresponding vertical profiles of length scale are shown for temperature in figure F.11. The regions are defined in table F.1. According to the ensemble method, the length scale is largest in the tropics and smallest in the data rich regions (for instance, in EUROPE and NORAMER areas). This result is in agreement with other studies concerning the effects of latitude on the length scales (e.g.

Lindzen and Fox-Rabinovitz (1989), Ingleby (2001)) and the effects of data density (Bouttier, 1994).

Table F.1: Regions used to study the heterogeneities of the background error statistics and to compute the forecast scores (as it will be shown in section F.5).

| Acronym | Region | latitudes | longitudes |
|---|---|---|---|
| NORD20 | Northern Hemisphere | 20°N - 90°N | |
| EUROPE | Europe (over land) | 30°N - 70°N | 10°W - 35°E |
| EURATL | Europe and East region of North Atlantic | 30°N - 70°N | 10°W - 35°E |
| NORAMER | North America (over land) | 25°N - 60°N | 145°W - 50°W |
| AMNORD | North America and surrounding ocean areas | 25°N - 60°N | 145°W - 50°W |
| ASIE | Asie | 25°N - 65°N | 60°E - 145°E |
| TROPIQ | tropics | 20°S - 20°N | |
| SUD20 | Southern Hemisphere | 20°S - 90°S | |
| AUS/NZ | Australia/New Zealand and surrounding ocean areas | 10°S - 55°S | 90°E - 160°E |

In contrast, between 950hPa and 500hPa, the largest $L^{\epsilon}_{NMC}$ values are found in the NORAMER region, rather than in the TROPIQ area. This is related to the fact that the differences between $L^{\epsilon}_{NMC}$ and $L^{\epsilon}_{ENS}$ are largest over NORAMER. This indicates that the background correlation overestimation is more pronounced over data rich areas. Moreover, contrarily to the ensemble method, it was found that in the middle troposphere, the $L^{\epsilon}_{NMC}$ value for temperature is larger over land than over sea (not shown). The comparison between $L^{\epsilon}_{ENS}$ and the length scales of $\epsilon^{i+4}_{a,NMC}$, $\epsilon^{i+6}_{b,NMC}$ and of the analysis increments (not shown), suggests that these differences are mainly caused by an inadequate representation of the analysis step (which relies on the accumulation and time evolution of several analysis increments during a 24h period, as shown in section F.2.3) in the NMC method.

It is also possible to diagnose the main direction and intensity of the local correlation anisotropies (see appendix). The figure F.12 illustrates the kind of features that can be highlighted. The correlations appear to be mainly elongated along a South-West/North-East direction in the $10-20°N$ latitude belt over Africa, and have a South-East/North-West elongation in the South Atlantic. It also worthwhile to mention that the areas of largest anisotropy appears to be associated to the maximum strength of jet stream. From figure F.12 it is also visible that the shape of the local correlation function (at the two selected points) is consistent with the direction of the main local anisotropy axis. This kind of local anisotropy diagnosis is likely to be particularly interesting in the future, to evaluate the properties of heterogeneous covariance formulations. For instance, these diagnostic equations appeared to be very informative when studying the properties of wavelets for the error covariances of the Aladin model (Deckmyn and Berre 2005).

### F.4.3 Vertical correlations

The latitudinal dependence of the vertical correlations of temperature is illustrated in figure F.13. Both ensemble and NMC methods indicate that the vertical correlations become narrower in tropics.

Moreover, it appears that the mid-tropospheric background errors are negatively correlated with levels near the tropopause. However, when the NMC method is used, the vertical correlation functions are broader, and the negative correlations near the tropopause have a larger amplitude. When using the ensemble method, a negative correlation between middle and lower troposphere is visible, while in the NMC method, the positive correlation extends from the mid-troposphere to the surface.

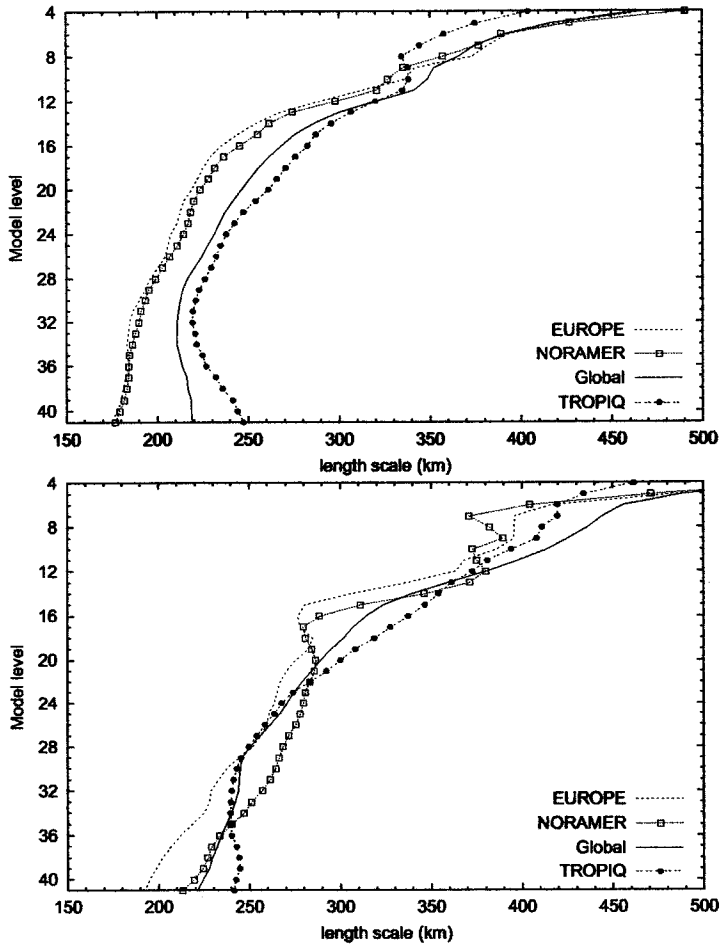


Figure F.11: Length scales of the background error of temperature, estimated by the ensemble method (top) and by the NMC method (bottom). The acronyms of the 3 areas are defined in table F.1.

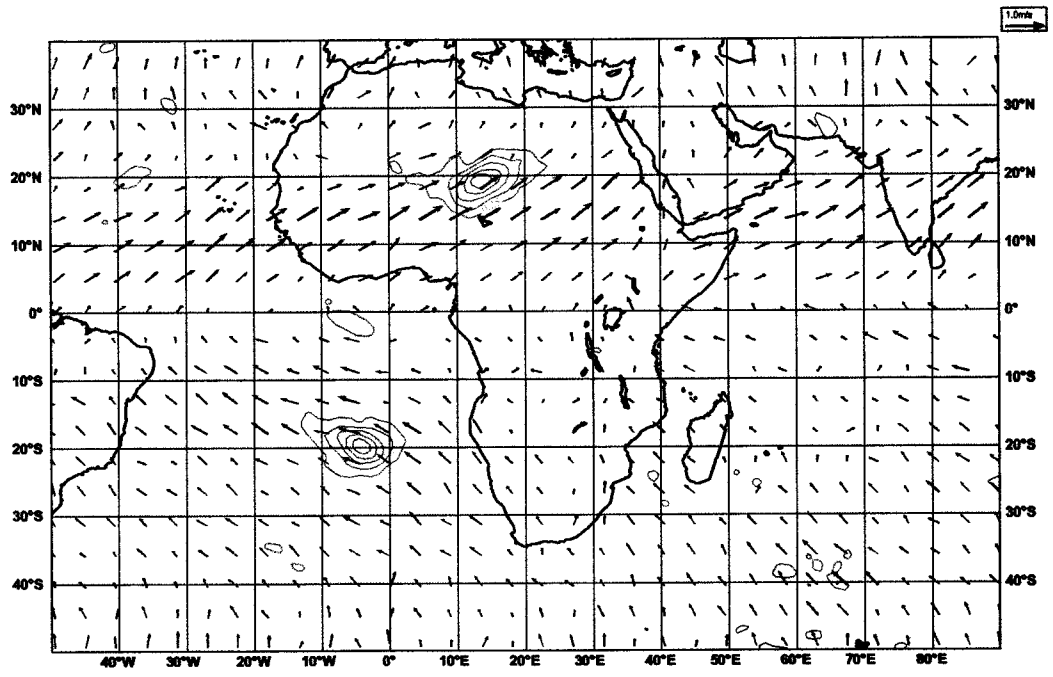


Figure F.12: Anisotropy vector of the local autocorrelation functions of the temperature background error at model level 21, overlapped with its autocorrelation function at (20°S, 5°W) and (20°N, 14°E). The isoline interval is 0.15 starting at 0.25. The arrows are arbitrarily oriented northwards, and their scaling is such that half of the vector length from the legend corresponds to an oblateness of 0.5.

The top panel of figure F.14 shows the vertical correlation functions of vorticity at level 21, estimated by the ensemble method, for the global average and for the EUROPE and TROPIQ areas. It is obvious that the vertical correlations are narrower in the TROPIQ area, as expected from other studies (Ingleby, 2001). Compared with the global average, the vertical correlation function over Europe is slightly broader in the short distances, and sharper in the long distances.

### F.4.4 Analysis versus background errors

The NMC and ensemble methods differ regarding the representation of the analysis step, as mentioned in section F.2. In this section, the corresponding estimates of the local standard-deviations are compared, for both analysis errors and background errors.

In the NMC method, the vectors $\epsilon^{i+4}_{a,NMC}, \epsilon^{i+6}_{b,NMC}$ may be seen as some estimates of the analysis errors and forecast errors, respectively (see section F.2.3: $\epsilon^{i+4}_{a,NMC}$ corresponds to some 24h-00h forecast differences and $\epsilon^{i+6}_{b,NMC}$ are the final 36h-12h forecast differences). The corresponding vertical profiles of standard deviation are represented in figure F.15, for the SUD20 (top panel) and for NORAMER (bottom panel) regions. These profiles can be also compared with the corresponding ensemble estimates, that are represented in figure F.16. These ensemble estimates correspond to the respective dispersions of the analyses and of the backgrounds of the ensemble.

For the NMC method, it appears that the maximum magnitude of both the analysis and the background errors is slightly larger over NORAMER than in SUD20 region. On the contrary, for the ensemble method, the analysis and background dispersions are larger in SUD20 than over NORAMER region. The ensemble results appear therefore to be more consistent with the expected effects of the data density contrasts (i.e. more observations over North America than in the Southern Hemisphere).

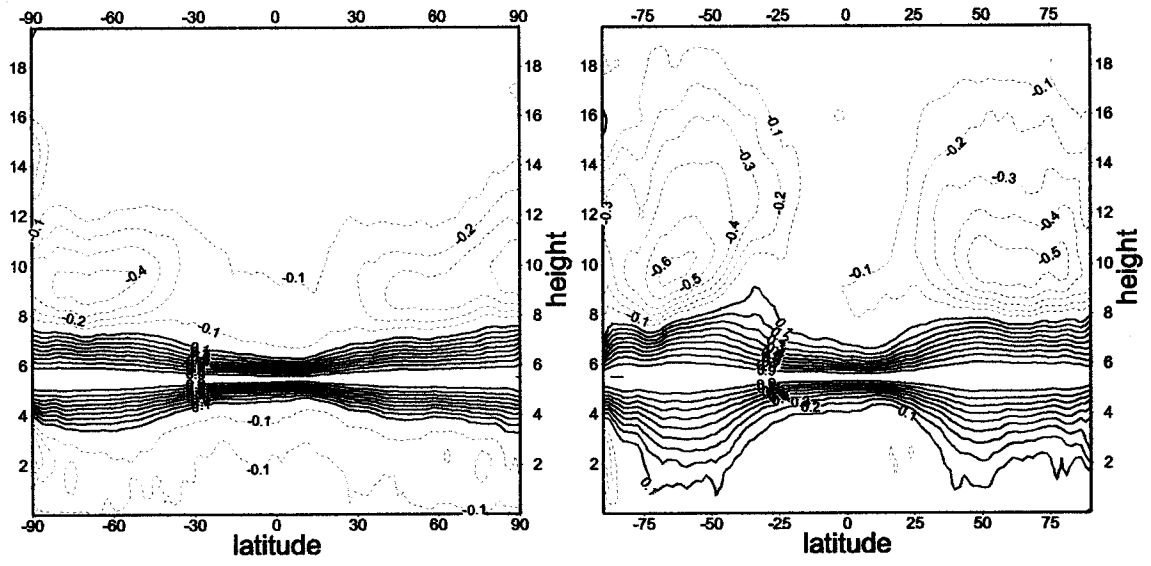


Figure F.13: North-South variation of the vertical correlations of temperature background error (at level 21, i.e. around 500 hPa), estimated by the ensemble method (left) and by the NMC method (right). Isoline spacing is 0.1.

For the ensemble method, it can be also noticed, e.g. around the jet level, that the difference between the analysis and background dispersions is relatively larger over NORAMER than over SUD20 region: as expected, the analysis effect is stronger in the data rich area than in the data poor area. Similarly, the background dispersion appears to be smaller at 18 UTC than at 12 UTC. This seems to be consistent with the larger amount of available observations at 12 UTC than at 06 UTC.

These results support therefore the idea that the analysis effects (on the error evolution) are more adequately represented in the ensemble method than in the NMC method.

## F.5 Impact on the forecast scores

### F.5.1 Experiments

Some assimilation and forecast experiments have been performed in order to investigate the impact of the ensemble statistics, compared with the statistics that were used in the operational configuration (as described in section F.2.4). The impact of both the ensemble global covariances and of the ensemble local standard deviations has been studied. In a similar way as for the operational local standard deviations, two different types of ensemble local standard deviations are involved. Firstly, a 3D map of vorticity standard deviations is used in the minimization. In other words, this map is a part of the **B** matrix which is involved in the cost function, and which determines the weights of the background in the analysis solution. Secondly, some 3D maps of standard deviations (for wind, temperature, and geopotential) are used in the quality control. These quality control maps are obtained from the "vorticity maps" and from the ensemble global covariances, by using a randomization technique (Andersson and Fisher 1998). These two maps are both truncated at T21, in order to provide some smooth geographical variations.

The list of experiments and their features are summarized in table F.2. The experiments were performed for the period between 5 February and 4 March 2002. Some experiments have been also performed over an October/November 2002 period and the main results were roughly similar.

In addition, some experiments have been performed in order to investigate the impact of

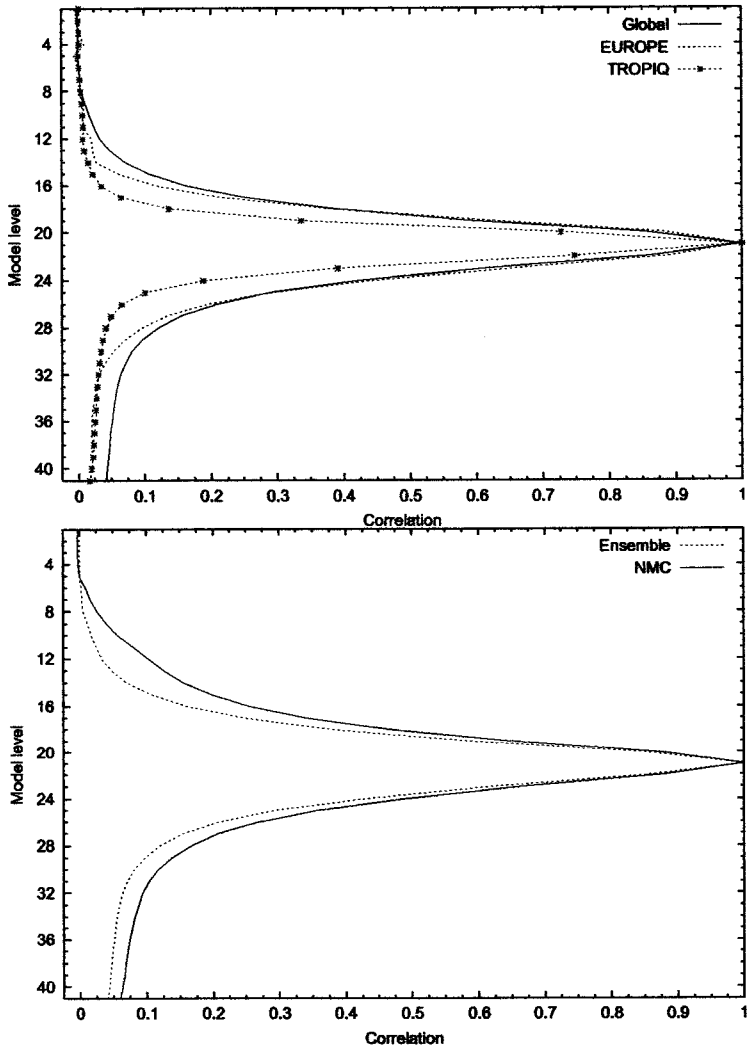


Figure F.14: Vertical correlations of the vorticity background error (at level 21, i.e. around 500 hPa). Top panel: ensemble estimations over different parts of the globe. Bottom panel: global estimates by the ensemble method and by the NMC method.

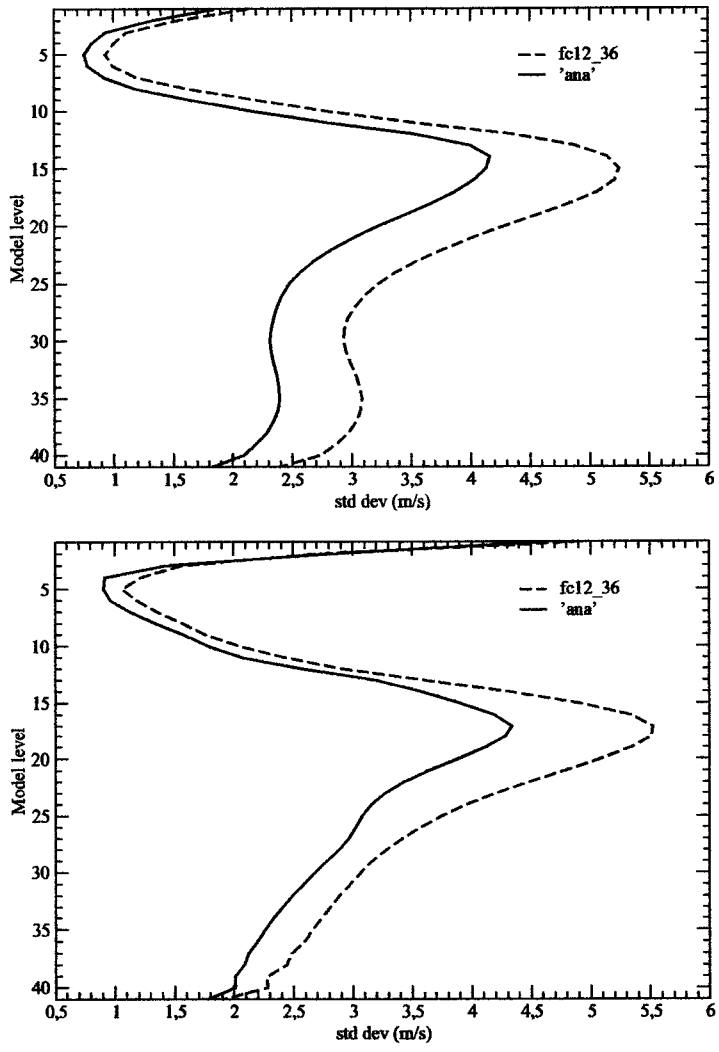


Figure F.15: Standard deviations of the analysis error (solid line) and of the background error (dashed line), for zonal wind, estimated by the NMC method. Top panel: SUD20. Bottom panel: NORAMER (see table F.1).

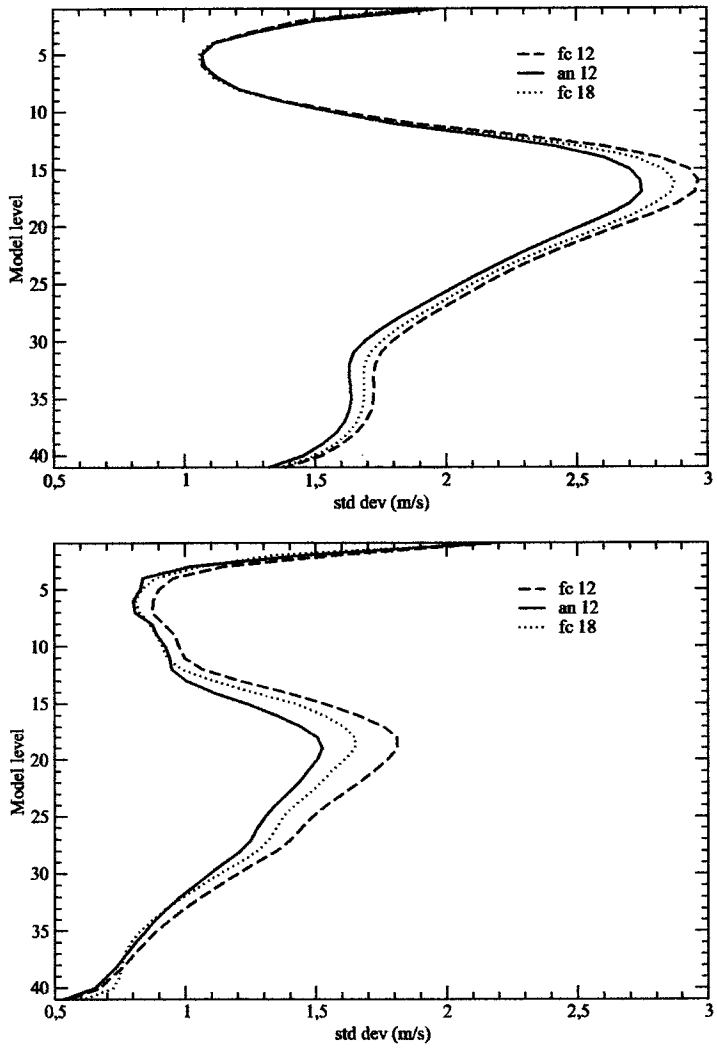


Figure F.16: Standard deviations of the analysis error at 12 UTC (solid line), and of the background error at 12UTC (dashed line) and at 18UTC (dotted line), for zonal wind, estimated by the ensemble method. Top panel: SUD20. Bottom panel: NORAMER (see table F.1).

diurnal cycle on the ensemble statistics. For this purpose, different global and local background error statistics have been used at each analysis time in the 4D-Var assimilation cycle. The impact of these changes was mostly neutral.

The forecast scores were calculated by comparing the forecasts with both ECMWF analysis and TEMP observations, for geopotential, wind, temperature and relative humidity. The main results were similar with both verification sources. The impact of the ensemble statistics were mostly neutral for humidity. The largest impact in the scores was found for geopotential and wind (in a similar way). Therefore, we will simply present the wind scores with respect to the ECMWF analysis.

Table F.2: Experiments. The following letters are used to refer to the different statistics: ENS = ensemble statistics, NMC = NMC method's global covariances, OLS = operational local standard deviations. The last two columns refer to the local standard deviations that are used in the minimization (min.) and in the quality control (q.c.), respectively.

| Experiment | Global covariances | Local std. dev. (min.) | Local std. dev. (q.c.) |
|---|---|---|---|
| OPE | NMC | OLS | OLS |
| EB0 | ENS | OLS | OLS |
| ER3 | ENS | ENS | OLS |
| GR3 | ENS | ENS | ENS |

### F.5.2 Impact of the ensemble global covariances

In the operational version of Arpège 4D-Var, the global NMC standard-deviations are multiplied by a scaling factor $R$, which is equal to $R = 0.9$.

As mentioned in section F.3.1, the ensemble estimates of the global standard deviations are much smaller than with the NMC method. Therefore, a scaling factor equal to $R = 1.5$ has been applied to the ensemble standard deviations, so that the resulting magnitudes are similar to the operational ones. In addition, some sensitivity studies have indicated that this choice of $R$ was close to optimal (the scores were slightly degraded or neutral, when decreasing or increasing $R$ by 10%).

The left panels of figure F.17 correspond to the impact of the ensemble global covariances, compared with the NMC global covariances. The impact appears to be clearly positive over the AMNORD and EURATL domains, with some increase of the impact amplitude when the forecast range increases. The scores are also generally positive over the tropics. In the other regions, the scores are slightly positive or neutral.

### F.5.3 Impact of the ensemble local standard deviations

The middle panels of figure F.17 correspond to the total impact of both the ensemble global covariances and the ensemble local standard deviations (of vorticity). The generally positive impact, which was visible in the left panels, appears to be enhanced due to the use of the ensemble local standard deviations. This is particularly noticeable over the ASIE and EURATL areas. This result suggests that the enhancement (by the ensemble method) of the standard deviation in these regions (see figure F.7) is relevant.

The right panels of figure F.17 correspond to the total impact of the three types of ensemble statistics, compared with the operational statistics: the third involved ensemble statistics are the local ensemble standard deviations, that are used in the quality control stage of the 4D-Var analysis. The addition of this third ingredient appears to strengthen the previous positive impacts. This is more obvious over the AMNORD, ASIE and EURATL areas.

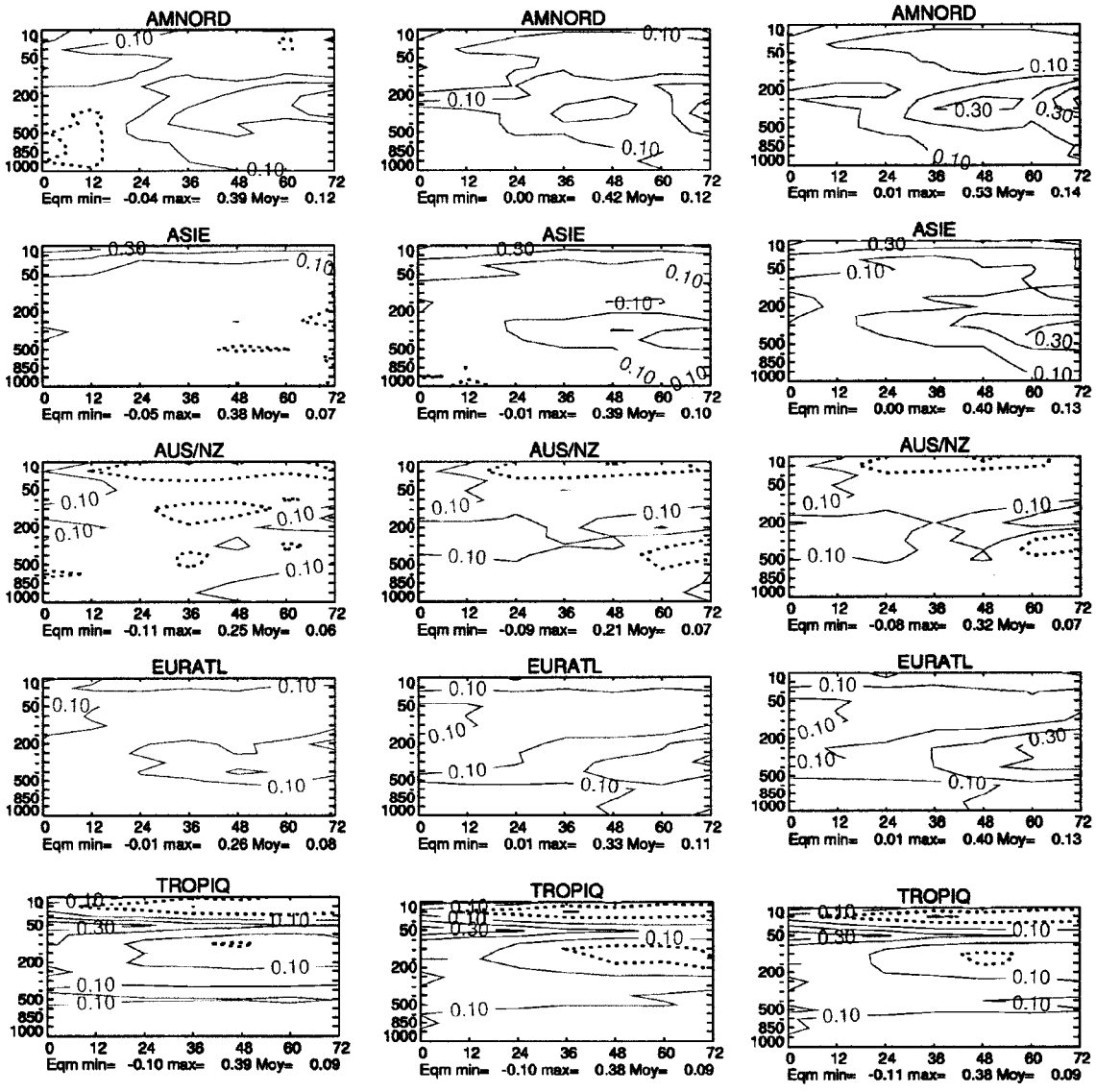


Figure F.17: Difference between the root mean square errors (RMSE, in m $s^{-1}$) of wind forecasts (as function of pressure levels and forecast range) of the operational experiment (OPE) and those of the following experiments: EB0 (left panels), ER3 (middle panels) and GR3 (right panels). Forecasts are verified against the ECMWF analysis. Solid (resp. dotted) isolines indicate ranges and levels, for which the scores of the involved experiment are better than (resp. similar to) the scores of the operational experiment. The isoline spacing is 0.10 m $s^{-1}$. The acronyms of the 5 areas are defined in table F.1.

All these results support the idea that there is some relevant and useful information in the ensemble estimates, for both the global and local statistics.

## F.6 Conclusions and perspectives

The estimation of the background error covariance matrix plays an important role in data assimilation. In the last years, the global covariances used in the Arpège 4D-Var, as in other NWP models, have been determined by using the NMC method. More recently, an interesting alternative technique based on an ensemble of analysis experiments was proposed by Houtekamer *et al.* (1996) and Fisher (2003). This technique has also been tested at Météo France, where it became operational in January 2004.

In the present paper we describe the results of the analysis ensemble experiments performed with the nonstretched version of the Arpège global model, in a perfect-model context. The ensemble estimates of the covariances were compared with those that were operational, and in particular with the results of the NMC method.

Globally, the ensemble approach appears to emphasize the relative contributions of the small scales to the background errors. Consequently, the correlation functions are sharper in ensemble method than in NMC method. These results are in accordance with the ones described by Fisher (2003), where these differences were considered to be caused by the involvement of longer forecast ranges in the NMC method. Nevertheless, the analysis increment is in fact one of the basic components of the NMC method and its spectrum was found to be much larger scale than the analysis dispersion spectrum (in the ensemble method). This suggests that the different representations of the analysis step (in the simulation of the error evolution) explain also an important part of the differences between the two methods.

Bouttier (1994) suggested that the analysis error variances estimated by the NMC method are expected to be underestimated in the data-poor areas, and overestimated in regions where the observations are dense and have a high quality. For these reasons, in the Arpège 4D-Var, the NMC method is not used to provide the spatial variations of the background error variance. The examination of the local covariances, and the comparison between the analysis errors and the background errors estimated by the NMC method, support the idea argued by Bouttier (1994).

In contrast with this, the local ensemble background error variances capture some interesting features, such as the contrasts between data-rich and data-sparse regions, and also the areas of large atmospheric variability (the ITCZ, for instance). Moreover, the analysis dispersion provided by the ensemble approach appears to give some relevant information about the analysis error, for instance concerning the influence of data density.

The local correlation structures were also examined. Generally, in NWP models, the computation limitations are an hindrance to the diagnosis and representation of the spatial variations of the horizontal correlations. An economical method was therefore presented to estimate the local correlation length scale. The corresponding estimates provided by the ensemble method show that the length scale is largest in the tropics and smallest in the data rich regions, as it would be expected from other studies (e.g. Lindzen and Fox-Rabinovitz (1989), Ingleby (2001), Bouttier (1994)). In addition, these estimates from the NMC method differ from the ensemble ones, mainly for temperature. In particular, the results suggest that the overestimation of the length scales by the NMC method is more pronounced over data rich areas, such as North America. These differences suggest that the expected data density effects on the correlation structures are not well represented by the NMC method.

Moreover, this local length scale diagnosis allows also to examine the correlation anisotropies. These informations about the heterogeneities and anisotropies can be useful to model spatially varying correlations in the **B** matrix, by wavelet techniques (Fisher (2003), Deckmyn and Berre (2005)) for instance.

Some impact studies were finally performed with the Arpège 4D-Var. The use of the ensemble global covariances (compared with the NMC global covariances) had a general positive impact

on the forecast quality. This positive impact was found to be enhanced when the ensemble local standard deviation maps (instead of the former operational ones) are used in the minimization and quality control stages. These results confirm the indications of the diagnostic studies: there is some relevant global and local information in the ensemble method, and this technique is in particular more appropriate than the NMC method.

Nevertheless, the ensemble approach applied here assumes that the forecast model is perfect. It is expected that this assumption causes an underestimation of the error variances. Therefore, in the near future we intend to introduce and test some model perturbations, in order to study the impact of the model error on the background error covariances.

The ensemble simulations presented in this paper have been done during a winter period. In the future, it would be also interesting to perform such experiments during other seasons. This would allow to examine the seasonal variations of e.g. the standard deviation maps (including for instance a detailed comparison with the results of the cycling algorithm that is described in Derber and Bouttier (1999)). More generally, the relevance of the ensemble local estimates suggests that there may be some interesting flow-dependent information to be extracted from the ensemble dispersion. In this perspective, it would be interesting to examine the potential of e.g. wavelets to extract these local flow-dependent signals.

*Acknowledgements.* The authors would like to thank M. Fisher for his help with the observation perturbation code. The authors also would like to thank F. Bouttier, F. Bouyssel, C. Fischer and G. Desroziers for the constructive comments on the manuscript. Margarida Belo Pereira is grateful to the ALATNET Grant HPRN-CT-1999-00057. ALATNET is supported by the TMR/IHP Program of the European Community, but the information provided here is responsibility only of the ALATNET team. Margarida Belo Pereira is also grateful to Météo France for partly supporting her stays in CNRM/GMAP.

## F.7 Appendix: Local length scales and anisotropy vectors

The background error variance of each variable and of its derivative can be used to estimate the horizontal correlation length scale. For the sake of simplicity, we present here the derivations for the one-dimensional case. For any simulated error $\epsilon$, the covariance of the derivative $\frac{\partial \epsilon}{\partial x}$ between two points can be expressed as:

$$\left\langle \frac{\partial \epsilon_1}{\partial x_1} \frac{\partial \epsilon_2}{\partial x_2} \right\rangle = \frac{\partial^2}{\partial x_1 \partial x_2} \left( \langle \epsilon_1 \epsilon_2 \rangle \right) \tag{F.5}$$

Here, $x_1$ and $x_2$ are the coordinates of the two points being correlated.

Furthermore, the background error covariance can be defined as a function of its autocorrelation function ($\rho$), and of its variances $\sigma(\epsilon_1)$ and $\sigma(\epsilon_2)$ at points $x_1$ and $x_2$, respectively:

$$\langle \epsilon_1 \epsilon_2 \rangle = \sigma(\epsilon_1)\sigma(\epsilon_2)\,\rho \tag{F.6}$$

So, the second derivative of $\langle \epsilon_1 \epsilon_2 \rangle$ is

$$\begin{aligned} \frac{\partial^2}{\partial x_1 \partial x_2} \left( \langle \epsilon_1 \epsilon_2 \rangle \right) = \frac{d\sigma(\epsilon_1)}{dx_1} \frac{d\sigma(\epsilon_2)}{dx_2} \rho + \sigma(\epsilon_1)\sigma(\epsilon_2) \frac{\partial^2 \rho}{\partial x_1 \partial x_2} + \\ + \sigma(\epsilon_1) \frac{d\sigma(\epsilon_2)}{dx_2} \frac{\partial \rho}{\partial x_1} + \sigma(\epsilon_2) \frac{d\sigma(\epsilon_1)}{dx_1} \frac{\partial \rho}{\partial x_2} \end{aligned} \tag{F.7}$$

Defining $\tilde{x} = x_2 - x_1$ allows to write that $\frac{\partial}{\partial x_1} = -\frac{d}{d\tilde{x}}$ and $\frac{\partial}{\partial x_2} = \frac{d}{d\tilde{x}}$. In addition, the first derivative of $\rho$ (it is supposed that $\rho$ is derivable) can be assumed to be equal to zero at $\tilde{x} = 0$. This assumption looks reasonable, knowing that the value of $\rho$ decreases when the separation distance increases from the origin ($\tilde{x} = 0$), and assuming that $\rho$ is continuous near its origin.

Replacing equation (F.5) into equation (F.7), and considering the case $\tilde{x} = 0$, provides the following expression for the local variance of $\frac{\partial \epsilon}{\partial x}$:

$$\sigma^2\left(\frac{\partial \epsilon}{\partial x}\right) = \left( \frac{d\sigma(\epsilon)}{dx} \right)^2 - \sigma^2(\epsilon) \left. \frac{d^2 \rho}{dx^2} \right|_{x=0}$$

Finally, when using the definition of length scale given by Daley (1991), for the one-dimensional case

$$L^2 = - \left. \frac{\rho(x)}{d^2\rho/dx^2} \right|_{x=0} \tag{F.8}$$

we obtain the correlation length scale for the background error, $\epsilon$, of any variable, as a function of the variances of $\epsilon$ and of $\frac{\partial \epsilon}{\partial x}$:

$$L_x^{\epsilon} = \sqrt{\frac{\sigma^2(\epsilon)}{\sigma^2\left(\frac{\partial \epsilon}{\partial x}\right) - \left(\frac{d\sigma(\epsilon)}{dx}\right)^2}}. \tag{F.9}$$

It is interesting to note that due to the Helmholtz's theorem, for streamfunction ($\psi$) the correlation length scale can be written as a function of the background error variances of $\psi$ and of the rotational part of the meridional wind ($v^{\psi} = \frac{\partial \psi}{\partial x}$):

$$L_x^{\psi} = \sqrt{\frac{\sigma^2(\psi)}{\sigma^2(v^{\psi}) - \left(\frac{d\sigma(\psi)}{dx}\right)^2}}. \tag{F.10}$$

Hollingsworth (1987) and Bouttier (1993) have studied the anisotropy of correlation, by computing the inertia matrix of the correlation function. In a similar way, the anisotropy of $\rho$ can be determined from the eigenvectors and eigenvalues of the following matrix $\mathcal{N}$:

$$\begin{pmatrix} 1/\mathcal{N}_{xx} & 1/\mathcal{N}_{xy} \\ 1/\mathcal{N}_{xy} & 1/\mathcal{N}_{yy} \end{pmatrix} \tag{F.11}$$

where $\mathcal{N}_{xx} = L_x^2$ and $\mathcal{N}_{yy} = L_y^2$ (where $L_x, L_y$ are the zonal and meridional length scales, that are defined in section F.4.2). The term of tilting of $\rho$ is determined by:

$$\mathcal{N}_{xy} = \frac{\sigma^2(\epsilon)}{\langle \frac{\partial \epsilon}{\partial x} \frac{\partial \epsilon}{\partial y} \rangle - \frac{\partial \sigma(\epsilon)}{\partial x} \frac{\partial \sigma(\epsilon)}{\partial y}} \tag{F.12}$$

for any error $\epsilon$. Here $\langle \frac{\partial \epsilon}{\partial x} \frac{\partial \epsilon}{\partial y} \rangle$ is the covariance between $\frac{\partial \epsilon}{\partial y}$ and $\frac{\partial \epsilon}{\partial x}$.

The oblateness ($\mathcal{O}$) of the correlation function is defined as $\mathcal{O} = 1 - \lambda_2/\lambda_1$, where $\lambda_1$ and $\lambda_2$ are the largest and the smallest eigenvalues, respectively. The anisotropy of $\rho$ is defined by a vector, which has the direction of the eigenvector of $\mathcal{N}$ corresponding to its smallest eigenvalue. In other words, the main anisotropy axis identifies the direction of the largest elongation of the correlation function. Moreover, the norm of the anisotropy vector is given by $\mathcal{O}$: therefore, a null vector is equivalent to an isotropic $\rho$. An oblateness equal to 0.5 means that the value of $\rho$ decreases with distance twice faster in one direction than in the perpendicular one.

## REFERENCES


Andersson, E., and M. Fisher, 1998: Background errors for observed quantities and their propagation in time. *Proceedings of the ECMWF workshop on "Diagnosis of data assimilation systems"*, Reading, UK, 1-4 November 1998, p.81-89.

Berre, L., 2000: Estimation of synoptic and mesoscale forecast error covariances in a limited area model. *Mon. Wea. Rev.*, **128**, 644-667.

Berre, L., S.E., Ştefănescu, and M. Belo Pereira, 2006: The representation of the analysis effect in three error simulation techniques. To appear in *Tellus*.

Bouttier, F., 1993: The dynamics of error covariances in a barotropic model. *Tellus*, **45A**, 408-423.

Bouttier, F., 1994: Sur la prévision de la qualité des prévisions météorologiques. Ph. D. dissertation, Université Paul Sabatier, 240 pp. [ Available from Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex, France. ]

Daley, R., 1991: *Atmospheric data analysis.* Cambridge University Press, Cambridge, UK, 460 pp.

Deckmyn, A. and L. Berre, 2005: A wavelet approach to representing background error covariances in a limited area model. *Mon. Wea. Rev.*, **133**, 1279-1294.

Derber, J., and F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195-221.

Desroziers, G., V. Mathiot, and F. Orain, 1995: A study of ARPEGE forecast error covariances. *Proceedings of the WMO second international symposium on assimilation of observations in meteorology and oceanography*, Tokyo, Japan, Vol I, 263-268.

Fisher, M., 2003: Background error covariance modelling. *Proceedings of the ECMWF Seminar on Recent developments in data assimilation for atmosphere and ocean, 8-12 September 2003*, 45-63.

Gauthier, P., C. Charette, L., Fillion, P., Koclas, and S. Laroche, 1999: Implementation of a 3D variational data assimilation system at the Canadian Meteorological Centre. Part I: The global analysis. *Atmosphere-Ocean*, **37**, 103-156.

Gustafsson, N., L. Berre, S., Hörnquist, X.-Y., Huang, M., Lindskog, B., Navascués, K. S., Mogensen, and S. Thorsteinsson, 2001: Three-dimensional variational data assimilation for a limited area model. Part I: General formulation and the background error constraint. *Tellus*, **53A**, 425-446.

Hollingsworth A., 1987: Objective analysis for numerical weather prediction. Short- and medium-range numerical weather prediction. *Proc. WMO/IUGG NWP Symp.*, Tokyo, Japan, Meteorological Society of Japan, 11-59.

Houtekamer, P.L., L. Lefaivre, J., Derome, H., Ritchie, and H.L. Mitchell, 1996: A System Simulation Approach to Ensemble Prediction. *Mon. Wea. Rev.*, **124**, 1225-1242.

Ingleby, N.B., 2001: The statistical structure of forecast errors and its representation in The Met. Office Global 3-D Variational Data Assimilation Scheme. *Q. J. Roy. Meteor. Soc.*, **127**, 209-231.

Lindzen, R.S. and M. Fox-Rabinovitz, 1989: Consistent Vertical and Horizontal Resolution. *Mon. Wea. Rev.*, **117**, 2575-2583.

Lorenc, A.C., S.P. Ballard, R.S., Bell, N.B., Ingleby, P.L.F., Andrews, D.M., Barker, J.R., Bray, A.M., Clayton, T., Dalby, D., Li, T.J., Payne, and F.W. Saunders, 2000: The Met. Office Global 3-Dimensional Variational Data Assimilation Scheme. *Q. J. Roy. Meteor. Soc.*, **126**, 2991-3012.

Mc Nally, A., 2000: Estimates of short-term forecast-temperature error correlations and the implications for radiance-data assimilation, *Q.J.R. Meteorol. Soc.*, **126**, 361-373.

Mitchell, H.L., P.L. Houtekamer, and G. Pellerin, 2002: Ensemble Size, Balance, and Model-Error Representation in an Ensemble Kalman Filter. *Mon. Wea. Rev.*, **130**, 2791-2808.

Parrish, D.F. and J.C. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747-1763.

Rabier, F., A. Mc Nally, E., Andersson, P., Courtier, P., Undén, J., Eyre, A., Hollingsworth, and F. Bouttier, 1998: The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part II: Structure functions. *Quart. J. Roy. Meteor. Soc.*, **124**, 1809-1829.

Rabier, F., H. Järvinen, E., Klinker, J.-F., Mahfouf, and A.J. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. Part I: Experimental results with simplified physics. *Q.J.R. Meteorol. Soc.*, **126**, 1143-1170.

Ştefănescu S. E., L. Berre and M. Belo Pereira, 2006: The evolution of dispersion spectra and the evaluation of model differences in an ensemble estimation of error statistics for a limited area model analysis. Submitted to *Mon. Wea. Rev.*.

Veersé, F. and J.-N Thépaut, 1998: Multiple-truncation incremental approach for four-dimensional variational data assimilation. *Q.J.R. Meteorol. Soc.*, **124**, 1889-1908.

Wu, W.-S., R. Purser, and D. Parrish, 2002: Three-Dimensional Variational Analysis with Spatially Inhomogeneous Covariances. *Mon. Wea. Rev.*, **130**, 2905-2916.

Žagar N., E. Anderss05-2916.

Žagar N., E. Andersson, and M. Fisher, 2004: Balanced tropical data assimilation based on a study of equatorial waves in ECMWF short-range forecast errors. *ECMWF Technical Memorandum* No. 437, 30pp.

# Appendix G

**The representation of the analysis effect in three error simulation techniques**
***(paper to appear in Tellus)***

L. Berre[1], Simona Ecaterina Ştefănescu[2] and M. Belo Pereira[3]
[1]*Météo-France - CNRM/GMAP, Toulouse, France*
[2]*National Meteorological Administration, SMDCA, Bucharest, Romania*
[3]*Instituto de Meteorologia - DVM/VPN, Lisboa, Portugal*

Abstract

Three error simulation techniques are compared formally, in particular regarding their representation of the analysis step. The associated results are moreover examined for the Aladin-France limited area model, which is coupled with the Arpège global model.

It is first shown that the analysis error equation involves the same operators as the analysis equation. This implies that the analysis ensemble approach is appropriate, as the analysis equation is used to transform the background and observation dispersions into the analysis dispersion.

By contrast, the standard NMC method relies essentially on the analysis increment equation, which contributes to a large extent to the excessive emphasis on the large scale structures. The so-called lagged NMC method is shown to be closely related to the Arpège/Aladin model differences.

The analysis ensemble approach gives error spectra that are intermediate between those of the two other methods. This is in agreement with the representation of the initial and lateral boundary uncertainties, in a way that is consistent with the influence of the analysis equation and with the short forecast ranges.

## G.1 Introduction

Data assimilation schemes, such as three-dimensional variational systems (3D-Var), combine observations and a background, which is for instance a six hour (6h) forecast. In such schemes, the background error covariances determine the filtering and the propagation of the observed information. Nevertheless, the estimation of these error covariances is not straightforward, e.g. because the truth is never exactly known.

In this study, a method based on an ensemble of analyses is compared to two other simulation techniques, which are two variants of the NMC (National Meteorological Center, nowadays named National Center for Environmental Prediction) method (the so-called "standard" and "lagged" versions of the NMC method, according to the terminology of Široká et al., 2003). These three methods will be compared formally, with respect to the equations of the error evolution. In particular, the representation of the analysis step will be examined. This will allow to highlight the strong role of the analysis equation in the exact error evolution and in the ensemble simulation. The contrast between this and the involvement of the analysis increment equation in the NMC method will be made explicit and discussed. Moreover, some experimental results will be presented for the Aladin limited area model (LAM) (Bubnová et al., 1993 or Radnóti et al., 1995), which is coupled with the Météo France global model Arpège.

Techniques such as the NMC method are used in many NWP centres, but the related formalism and approximations are not often explicit. The long forecast ranges of the NMC method (for example 12 and 36 hours) are a well known drawback. In the present paper, the analysis step representation will be shown to be another critical component in the error simulation.

Some efforts to derive the implicit formalisms have nevertheless been done in the past. For instance, the formalism of the (standard) NMC method is made explicit and discussed in Bouttier (1994). The equations are presented in matrix form, and they concern essentially the covariance matrix equation of the first analysis increment. We propose therefore to explicit the NMC formalism in vector form (which facilitates e.g. the comparison with the ensemble formalism), and also to extend the discussion to the addition of several analysis increments (which is usually involved).

The results of the method based on an ensemble of analyses are described in Houtekamer et al. (1996) and Fisher (2003). A description of the corresponding formalism is presented in Žagar et al.(2004). The formalism of this approach will be established here in a partly different way, in the sense that it will be based on the usual BLUE ("best linear unbiased estimator") equation. Firstly, this facilitates the comparison with the NMC method. Secondly, this will show that the analysis equation can be applied to the error fields, even if the prescribed statistics and actual observation operator are not exact.

The formalism of the lagged NMC method will also be made explicit and discussed. It will be shown that this approach relies strongly on the Arpège/Aladin model differences (without a proper representation of the initial and lateral boundary uncertainties).

The paper is organized as follows. In section G.2, the equations that govern the evolution of the model state errors are derived, in particular to highlight the role of the analysis equation. Section G.3 describes the corresponding equations for the evolution of the model state differences in the ensemble error simulation method. A similar involvement of the analysis equation is in particular pointed out. Section G.4 deals with the implicit formalism of the standard NMC method. This allows to exhibit the strong role of the analysis increment equation. The associated formal expectations are then examined experimentally in section G.5, by comparing the Aladin-France results of the ensemble and standard NMC methods. In section G.6, the link between the lagged NMC method and the Arpège/Aladin model differences is examined, and the variance spectra of the three simulation methods are compared. Conclusions and perspectives are outlined in section G.7.

## G.2 The evolution of the model state errors

The equations that govern the evolution of the model state errors (i.e. the analysis errors and the background errors) will be derived. This will allow later on (sections G.3 and G.4) to explain the different representations of the analysis step in the ensemble and standard NMC methods. It will be shown that this analysis step difference contributes much to the experimental differences between the ensemble and standard NMC techniques in the Arpège/Aladin system (section G.5).

For simplicity, in the forecast and analysis steps, the operators will be considered to be linear. Similar equations can be derived in a non linear framework, for example by introducing some Taylor developments (Žagar et al., 2004).

### G.2.1 The forecast step

At time $t_i$, the analysis $x_a^i$ is an estimate of the true atmospheric state $x_*^i$, with uncertainties that correspond to the analysis error $e_a^i = x_a^i - x_*^i$.

The 6h forecast field $x_b^{i+1}$, that is valid at time $t_{i+1} = t_i + 6h$, is then obtained from this initial condition $x_a^i$, by integrating in time the forecast model, according to: $x_b^{i+1} = Mx_a^i$, where $M$ is the operator that corresponds to the six-hour evolution which is provided by the forecast model. The true fields will evolve according to: $x_*^{i+1} = M_* x_*^i$, where $M_*$ is the exact 6h forecast operator. This gives the following expression for the background errors $e_b^{i+1} = x_b^{i+1} - x_*^{i+1}$:

$$e_b^{i+1} = Me_a^i + e_m^{i+1} \tag{G.1}$$

where $e_m^{i+1}$ is the six-hour accumulated model error (see e.g. Daley, 1991, page 376) $e_m^{i+1} = (M - M_*)\, x_*^i$.

### G.2.2 The analysis step

The 6h forecast field $x_b^{i+1}$ will then be used as a background for the analysis at time $t_{i+1}$. The analysis equation is the equation that transforms the background and observation vectors into the analysis vector:

$$x_a^{i+1} = x_b^{i+1} + \mathbf{K}(y^{i+1} - Hx_b^{i+1}) \tag{G.2}$$

where $y^{i+1}$ is the observation vector at time $t_{i+1}$, $H$ is the observation operator, and $\mathbf{K}$ is the classical gain matrix, namely $\mathbf{K} = \mathbf{B}H^T(H\mathbf{B}H^T + \mathbf{R})^{-1}$. $\mathbf{B}$ and $\mathbf{R}$ are the specified spatial covariance matrices of the background errors and of the observation errors respectively, and the exponent $^T$ is a notation for the adjoint operator.

It is then possible to consider the application of the analysis equation to $x_*^{i+1}$ as a background field, and to $y_{*,H}^{i+1} = Hx_*^{i+1}$ as an observation vector. The vector $y_{*,H}^{i+1}$ is the projection of this true field onto the observation space, that is provided by the observation operator $H$. As the definition of $y_{*,H}^{i+1}$ implies that $y_{*,H}^{i+1} - Hx_*^{i+1} = 0$, it appears that the analysis equation applies well to the previous two fields:

$$x_*^{i+1} = x_*^{i+1} + \mathbf{K}(y_{*,H}^{i+1} - Hx_*^{i+1}) \tag{G.3}$$

The result $\mathbf{K}(0) = 0$ is valid as $\mathbf{K}$ is linear. More generally, this is also the case for a non linear analysis, if the analysis does not modify the background field, when the observations are exactly consistent with the background (i.e. when the observation vector is equal to $Hx_b^{i+1}$).

By calculating the difference between the last two equations, it becomes obvious that the analysis equation is also the equation that transforms the background and observation errors into the analysis errors (note that the related operators are the same as in equation (G.2)):

$$e_a^{i+1} = e_b^{i+1} + \mathbf{K}(e_o^{i+1} - He_b^{i+1}) \tag{G.4}$$

where $e_o^{i+1} = y^{i+1} - y_{*,H}^{i+1} = y^{i+1} - Hx_*^{i+1}$ is the observation error (which is itself the sum of the measurement error $e_{om}^{i+1} = y^{i+1} - y_*^{i+1}$, and of the representativeness error $e_{or}^{i+1} = y_*^{i+1} - y_{*,H}^{i+1} = y_*^{i+1} - Hx_*^{i+1}$, where $y_*^{i+1}$ is the exact atmospheric state in the observation space).

It may be interesting to notice also that the derivation of the analysis error expression does not require the assumption that the analysis is perfect: in particular, the gain matrix $\mathbf{K}$ may be suboptimal (i.e. the specified error covariance matrices $\mathbf{B}$ and $\mathbf{R}$ can be different from the exact ones), and the observation operator may be erroneous.

## G.3 The ensemble simulation method

The ensemble method that has been used e.g. by Houtekamer et al. (1996) and by Fisher (2003) consists in simulating the evolution of the model state errors. This is realized by introducing and evolving some perturbations, which have similar statistical features as the error contributions.

This approach may be described formally when considering the difference between two members of the ensemble, which will be referred to by the indices $k$ and $l$.

### G.3.1 The forecast step

At time $t_i$, two different analyses $x_{a,k}^i, x_{a,l}^i$ are available. Their difference is equal to $\varepsilon_a^i = x_{a,k}^i - x_{a,l}^i$.

Two forecast integrations are performed from these two initial states. Moreover, some model perturbations $\delta_{m,k}^{i+1}, \delta_{m,l}^{i+1}$ may be added, for instance at the end of the 6h forecasts. This provides two 6h forecast fields: $x_{b,k}^{i+1} = Mx_{a,k}^i + \delta_{m,k}^{i+1}$ and $x_{b,l}^{i+1} = Mx_{a,l}^i + \delta_{m,l}^{i+1}$. The difference between these two backgrounds, $\varepsilon_b^{i+1} = x_{b,k}^{i+1} - x_{b,l}^{i+1}$, reads:

$$\varepsilon_b^{i+1} = M\varepsilon_a^i + \varepsilon_m^{i+1} \tag{G.5}$$

with $\varepsilon_m^{i+1} = \delta_{m,k}^{i+1} - \delta_{m,l}^{i+1}$.

### G.3.2 The analysis step

The two different 6h forecast fields may then be combined with two different sets of observations $y_k^{i+1}, y_l^{i+1}$, in order to provide two different analyses at time $t_{i+1}$:

$$x_{a,k}^{i+1} = x_{b,k}^{i+1} + \mathbf{K}(y_k^{i+1} - Hx_{b,k}^{i+1})$$

$$x_{a,l}^{i+1} = x_{b,l}^{i+1} + \mathbf{K}(y_l^{i+1} - Hx_{b,l}^{i+1})$$

By calculating the difference between the last two equations, it appears that the analysis equation is the very same equation that transforms the background and observation differences into the analysis differences:

$$\varepsilon_a^{i+1} = \varepsilon_b^{i+1} + \mathbf{K}(\varepsilon_o^{i+1} - H\varepsilon_b^{i+1}) \tag{G.6}$$

where $\varepsilon_o^{i+1} = y_k^{i+1} - y_l^{i+1}$ is the vector of the observation differences. In practice, the two different sets of observations are obtained by adding two different perturbation vectors $\delta_{o,k}^{i+1}, \delta_{o,l}^{i+1}$ to the real observation vector $y^{i+1}$: $y_k^{i+1} = y^{i+1} + \delta_{o,k}^{i+1}$ and $y_l^{i+1} = y^{i+1} + \delta_{o,l}^{i+1}$.

The comparison between the equation pairs (G.1), (G.4) and (G.5), (G.6) is interesting: it indicates that the evolution processes and equations, that affect the ensemble difference fields, are the same as those of the true error fields.

In addition to this representation of the 6h forecast and analysis effects, some other requirements are involved, in order to achieve a realistic error covariance estimation. This involves the perturbation covariances of the observations and of the model. This issue and its implications on the error covariance estimations are evoked in the appendix.

## G.4 The standard NMC method

The NMC method (Parrish and Derber, 1992) computes differences between forecasts that are valid at the same time, but for different ranges, such as 12 and 36 hours. The 36h forecast may itself be seen as another 12h forecast, whose initial condition is a 24h forecast, instead of the operational analysis. This means that the initial differences between these two 12h forecasts correspond to the effect of the four successive steps of analysis and 6h forecast, which occur during the 24h period.

We propose to write formally the related equations, in order to compare them with the evolution equations of the errors and of the ensemble differences.

### G.4.1 The first assimilation cycle

At time $t_i$ occurring 30h before the verification time, there will be a first analysis step: the corresponding analysis increment $dx^i = x_a^i - x_b^i$ is the first difference that is introduced between the two integrations that will be compared.

This first analysis increment may be seen as an analysis perturbation $\varepsilon_a^i$, which is supposed to be an estimate of the analysis error $e_a^i$:

$$\begin{aligned} \varepsilon_a^i &= dx^i \\ &= \mathbf{K}(y^i - Hx_b^i) \\ &= \mathbf{K}(e_o^i - He_b^i) \end{aligned}$$

which allows to distinguish the respective contributions of the background and observation errors:

$$\varepsilon_a^i = -\mathbf{K}H\ e_b^i + \mathbf{K}\ e_o^i$$

This can be compared to the exact analysis error equation (according to equation (G.4)):

$$e_a^i = (I - \mathbf{K}H)\ e_b^i + \mathbf{K}\ e_o^i$$

and also to the analysis dispersion equation (following equation (G.6)):

$$\varepsilon_a^i = (I - \mathbf{K}H)\ \varepsilon_b^i + \mathbf{K}\ \varepsilon_o^i$$

The matrices $I - \mathbf{K}H$ and $\mathbf{K}$ define the respective weights of the background and observation errors, in the analysis error equation. In the NMC method, the background error weight $I - \mathbf{K}H$ is approximated by $-\mathbf{K}H$.

As discussed by Bouttier (1994), this may be a reasonable approximation if $\mathbf{K}H \sim I/2$. This corresponds to the case where the observations are very dense (i.e. $H \sim I$), and the two error covariances are approximately the same ($\mathbf{R} \sim H\mathbf{B}H^T \sim \mathbf{B}$). The latter condition means that the observations are considered to have a similar quality and similar spatial error structures (i.e. similar error variances and correlations) as the background.

Nevertheless, in regions where data density is poor or where observations have a poor quality, the amplitude of the analysis increments is likely to be small, while the amplitude of the analysis error is large.

Moreover, the observation errors are usually much less correlated spatially than the background errors: the observation error spectrum tends to be white, while the background error spectrum is typically red. As discussed e.g. by Daley (1991, section 4.5), this implies that the operator $\mathbf{K}H$ tends to act as a low-pass filter. Conversely, the operator $I - \mathbf{K}H$ is expected to act as a high-pass filter.

This implies that the analysis increment is likely to be of a larger scale than the analysis error. In other words, the analysis error correlations are expected to be overestimated by the NMC method.

The NMC analysis perturbation $\varepsilon_a^i$ will evolve according to the forecast model during six hours, which will provide a NMC background perturbation $\varepsilon_b^{i+1}$:

$$\begin{aligned} \varepsilon_b^{i+1} &= M\ \varepsilon_a^i \\ &= M\ dx^i \end{aligned}$$

## G.4.2 The next three assimilation cycles

A second analysis step, which introduces some additional differences, will occur 24 hours before the verification time. This amounts to considering that a new analysis increment will be added to the background perturbation (namely $\varepsilon_b^{i+1} = M\varepsilon_a^i = Mdx^i$), in order to produce the following analysis perturbation:

$$\begin{aligned} \varepsilon_a^{i+1} &= Mdx^i + dx^{i+1} \\ &= \varepsilon_b^{i+1} + \mathbf{K}(e_o^{i+1} - He_b^{i+1}) \end{aligned}$$

This last equation indicates that, in the evolution of the NMC perturbations, the representation of the analysis effect consists in *adding the analysis increment to some earlier increments* (for instance $\varepsilon_b^{i+1} = Mdx^i$ in the current case). This differs from the ensemble method, for which the representation of the analysis effect consists in *applying the analysis equation to the perturbations* $\varepsilon_b, \varepsilon_o$.

After the fourth analysis, the final NMC analysis perturbation is as follows:

$$\begin{aligned} \varepsilon_a^{i+3} &= \varepsilon_b^{i+3} + dx^{i+3} \\ &= M^3dx^i + M^2dx^{i+1} + Mdx^{i+2} + dx^{i+3} \end{aligned}$$

where $M^j$ is the operator that corresponds to the forecast evolution during a period of $j \times 6$ hours. $\varepsilon_a^{i+3}$ corresponds to the 00h-24h differences.

The final 12 hour evolution of the fourth analysis perturbation reads:

$$\varepsilon_b^{i+5} = M^2\varepsilon_a^{i+3}$$

## G.4.3 The three specific components of the NMC method

In order to interprete the observed differences between the ensemble and NMC methods, one may therefore compare the following two synthetic equations. They summarize how the forecast differences (e.g. at time $i+5$) are obtained, respectively in the ensemble and NMC methods:

$$\varepsilon_b^{i+5} = M\ \varepsilon_a^{i+4}$$

$$\varepsilon_b^{i+5} = M^2\ (\sum_{j=1}^{4} M^{4-j}\ dx^{i+j-1}\ )$$

These two equations allow to identify the main three specific components of the NMC method, compared to the ensemble method:

- the involvement of longer forecast ranges (see the occurrence of the matrices $M^2, M^{4-j}$ instead of the 6h matrix $M$);
- the accumulation of several increments (see the occurrence of the operator $\sum$);
- the involvement of analysis increments $dx$, instead of analysis differences $\varepsilon_a$.

The respective contributions of these three components will be diagnosed experimentally in the next section.

## G.5 Analysis ensemble against the standard NMC method: experiments

In this section, the results will be focused on the differences regarding the correlations and the contributions of different scales. This will illustrate the importance of the analysis step differences. The total variances will be presented later on in section G.6.3.

### G.5.1 The Arpège/Aladin experiments

The results of the standard NMC method are compared to the results of the ensemble method over a 48 days period (04/02/2002-23/03/2002), using fields valid at 12 UTC. The Aladin ensemble data set has been obtained in two steps. Firstly, an ensemble of Arpège global assimilation cycles has been performed in a perfect model framework (in accordance with the formal description in section G.3, and with $\varepsilon_m = 0$). The second step is the production of the ensemble of Aladin 6h forecasts, with initial conditions and boundary conditions provided by the Arpège ensemble.

The Arpège assimilation system is a four-dimensional variational data assimilation (4D-Var) scheme (Veersé and Thépaut, 1998). The Arpège experiments have been done with its stretched version (i.e. the resolution is not uniform over the globe). The stretching factor is equal to 3.5, the truncation is T298, and there are 41 vertical levels. The Aladin model was integrated over the Aladin-France integration domain. The main geometrical characteristics for the Aladin-France integration domain are as follows: there are 41 vertical levels, the domain is square with lengths equal to $L_x$=$L_y$=2850 km, and the number of grid points in each direction is $J$=$K$=300; the grid resolution is thus $\delta x$=$\delta y$=9.5 km and the spectral truncation corresponds to $M$=$N$=149 ($M, N$ being the maximum wave numbers in each direction).

### G.5.2 Comparison between the two final statistics

The respective error correlation spectra and correlation length scales, provided by the ensemble and (standard) NMC methods, are shown in figure G.1. The correlation spectra are the variance spectra normalized by the total variance: they allow to compare the respective contributions of the different scales to the shape of the correlation functions.

The ensemble approach appears to provide error spectra that are shifted towards the small scales. As expected, this leads to smaller correlation length scales. The situation is nevertheless the opposite for the first 10 to 15 highest levels (near the tropopause and in the stratosphere). In particular, for temperature in the stratosphere, the ensemble method does not show the (unexplained) length scale decrease of the NMC method.

Moreover, the length scale differences are more pronounced for a large scale variable such as temperature, than for a small scale variable such as divergence. This is related to the larger differences in the large scale part of the correlation spectra.

The larger emphasis on the small scale contributions (by the ensemble method) is consistent with the results of Fisher (2003) over the globe, in which the NMC overestimation of the large scale contributions was attributed to the long forecast ranges.

On the other hand, another explanation of the observed discrepancies between the ensemble and NMC methods could be the difference in the representation of the analysis effect.

### G.5.3 Diagnosis of the three elementary contributions in the NMC method

The section G.4 has shown that the NMC method involves three main components: the analysis increments, their accumulation, and their time evolution.

In order to diagnose the respective influences of these three ingredients, some correlation spectra have thus been examined at different stages of the two simulation techniques for surface pressure (figure G.3). The uninitialized Arpège fields (projected onto the Aladin grid) have been used for the comparison, in order to concentrate on the specific effects of the different steps of the NMC method. Moreover, it may be noticed that the typical NMC/ensemble contrast for the Aladin spectra (top left panel of figure G.1) matches well with the contrast for the corresponding Arpège spectra (figure G.2).

The top left panel of figure G.3 allows to compare the spectrum of the "final" 36h-12h differences with the spectrum of the "initial" 24h-00h differences. The difference between these two spectra corresponds to the 12h evolution of the initial differences. This 12h evolution implies a slight increase of the relative contribution of the largest scales (at the expense of the contribution of the smaller scales).

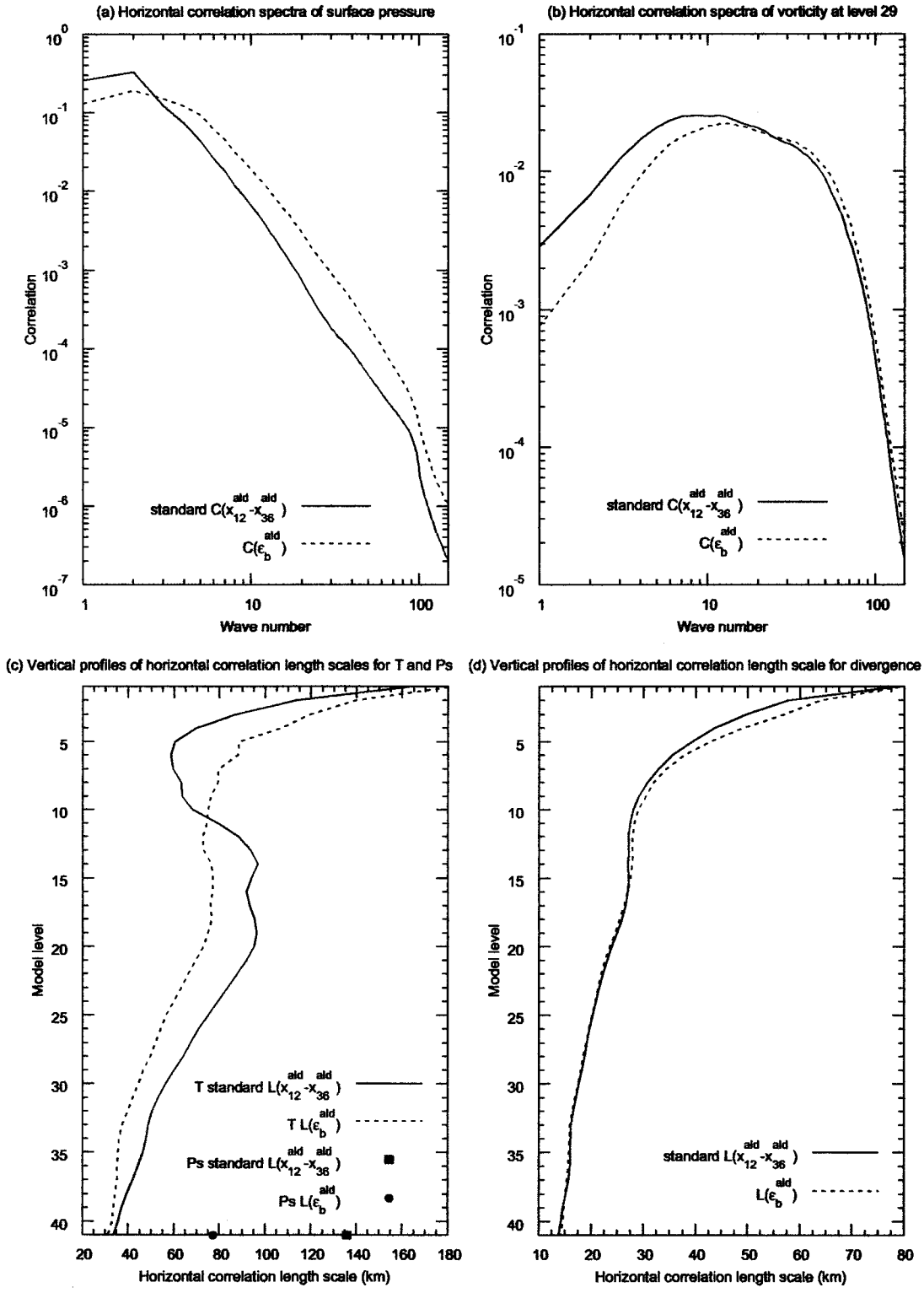


Figure G.1: The error correlation spectra (top panels) and the correlation length scales (bottom panels), for the Aladin standard NMC differences (full lines), and for the Aladin background ensemble differences (dotted lines). The model level 29 is located around the 775 hPa pressure level.

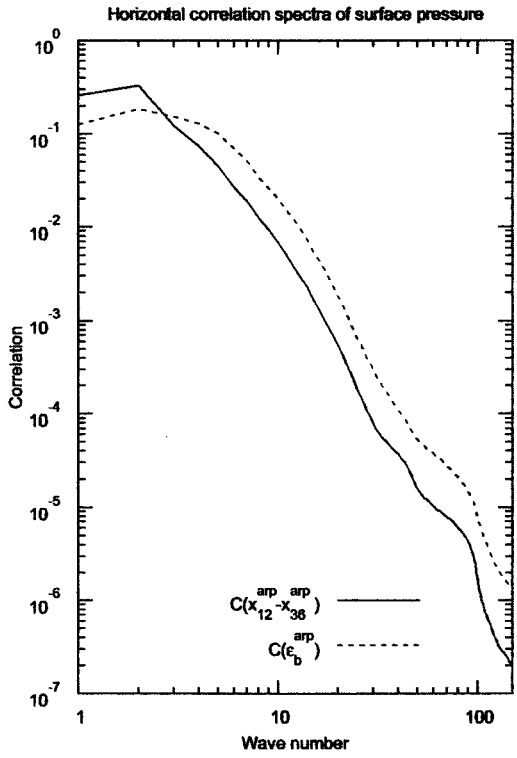


Figure G.2: The Arpège NMC correlation spectrum (full line) compared to the Arpège background ensemble correlation spectrum (dotted line).

The "initial" 24h-00h differences (which were described as the NMC analysis perturbations $\varepsilon_a^{i+3}$ in section G.4.2) are themselves the result of the accumulation of four successive analysis increments, of which the first three evolve in time:

$$\begin{aligned} \varepsilon_a^{i+3} &= \sum_{j=1}^{4} M^{4-j} dx^{i+j-1} \\ &= M^3 dx^i + M^2 dx^{i+1} + M dx^{i+2} + dx^{i+3} \end{aligned}$$

It is therefore interesting to diagnose the effect of the successive forecast evolutions of the first three analysis increments. This has been done by plotting the spectrum of the sum of the four analysis increments:

$$\sum_{j=1}^{4} dx^{i+j-1} = dx^i + dx^{i+1} + dx^{i+2} + dx^{i+3}$$

which amounts to replacing the forecast evolution matrices $M^{4-j}$ by the identity matrix.

The spectrum of the 24h-00h differences and the spectrum of the sum of the four analysis increments are presented in the top right panel of figure G.3. It appears that the evolutions of the (first three) increments contribute mostly to a slight enhancement of some intermediate large scales, corresponding to wave numbers between 4 and 20.

It is also possible to compare the spectrum of the increment sum $\sum_{j=1}^{4} dx^{i+j-1}$ with the specific spectrum of each increment $dx^{i+j-1}$. It appears (not shown) that the four increment spectra are very similar to each other, and that they are also close to the spectrum of the increment sum. Nevertheless, it may be mentioned that there is a slight enhancement of the contribution of the largest scales when the four increments are accumulated.

The bottom left panel of figure G.3 allows to visualize the total effect of the forecast evolutions and of the analysis increment accumulation. This panel represents indeed the respective spectra of the 12h-36h

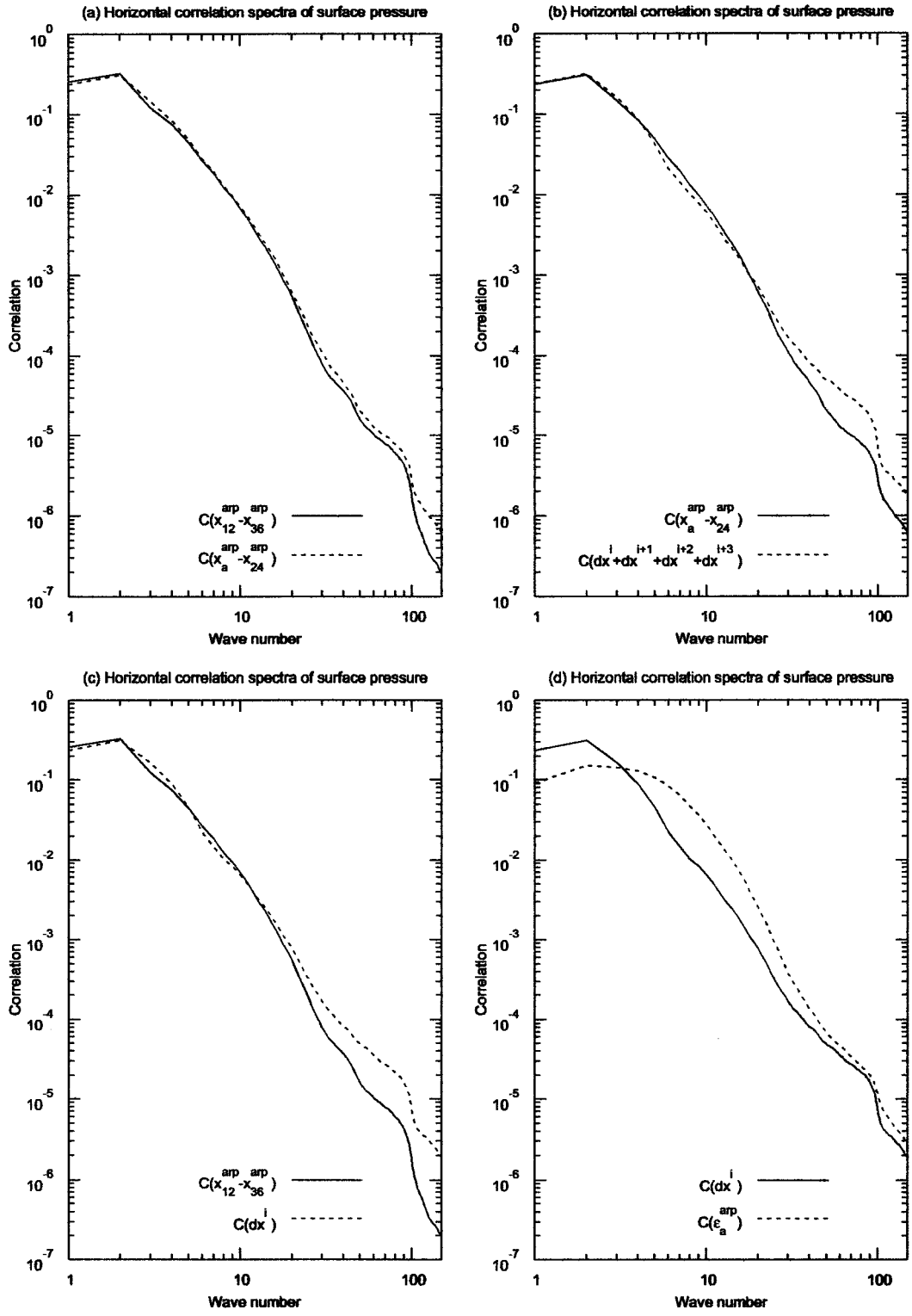


Figure G.3: The error correlation spectra of the following Arpège fields: (a) the 12h-36h NMC differences (full line) and the 00h-24h NMC differences (dotted line); (b) the 00h-24h NMC differences (full line) and the sum of the four (NMC) analysis increments (dotted line); (c) the 12h-36h NMC differences (full line) and the first (NMC) analysis increment (dotted line); (d) the first (NMC) analysis increment (full line) and the ensemble analysis differences (dotted line).

forecast differences, and of the first analysis increment $dx^i$. As expected, the forecast evolutions and the increment accumulation imply an increase of the large scale contributions globally, compared to the small scale contributions. On the other hand, the relative contributions of some small wave numbers (such as 3 and 4) are slightly reduced, and the global differences between the two spectra are less strong than in figure G.2.

It appears then interesting to compare (bottom right panel of figure G.3) the analysis ensemble spectrum with the first analysis increment spectrum $dx^i$ (of the NMC method). The discrepancy between these two curves corresponds to the influence of the differences (between the two simulation techniques) in the analysis effect representation.

The increment spectrum appears to be of a much larger scale than the analysis ensemble spectrum. This is consistent with the formal expectation that the correlations of the analysis increment are an overestimation of the analysis error correlations (see section G.4.1).

The first step of the NMC method, which relies on the first analysis increment, thus already emphasizes much more the large scale features than the ensemble method. Now, the contrast between the two related spectra is similar to what is observed in figure G.2, which confirms the importance of the analysis representation difference between the two simulation techniques.

To summarize, firstly, it appears that all the three NMC specific ingredients contribute to the increased emphasis on the large scales in the NMC method. Secondly, the different representations of the analysis step appear to be the strongest contribution to the scale differences between the two simulation techniques.

### G.5.4 The evolution of spectra during an analysis step

The influence of the analysis effect representation can be confirmed by examining what happens during the second analysis step of the NMC method. The left panel of figure G.4 shows the variance spectrum of the first increment after its six-hour evolution, namely $M dx^i$, and its modification when adding the second analysis increment $dx^{i+1}$. It appears that the addition of this second increment implies a general increase of the variances (which is somewhat stronger in the large scales than in the intermediate scales).

There is a strong discrepancy between this variance increase and the modification of the variance spectrum of the background dispersion during the analysis step (in the ensemble method) (right panel of figure G.4). The analysis contributes to a decrease of the large scale dispersion, which implies a reduction of the large scale contributions.

The latter result is consistent with the expectation that the analysis process uses the observations to reduce the amplitude of the large scale part of the background errors.

In other words, the large scale NMC/ensemble differences are related to the following different evolution processes: the NMC method is strongly based on the analysis increments (and on their accumulation and evolution), while the ensemble method is rather simulating the error reduction during the analysis step.

## G.6 The lagged NMC method

### G.6.1 Description of the method

A variant of the standard NMC method has been used by Široká et al. (2003), in order to derive background error covariances for the Aladin model. This variant has been called the lagged NMC method.

The difference from the standard NMC method concerns the lateral boundary conditions (LBCs) and the initial condition (IC) of the "fresh" 12h forecast integration. Instead of the usual operational ("fresh") fields, the LBCs and the IC of the 12h forecast correspond now to some fields that are derived from the "old" 36h forecast integration.

More precisely, the LBCs of the 12h forecast are taken as being identical to those that are used during the last 12 hours of integration of the 36h forecast. Moreover, the IC of the 12h forecast is now taken as being the coupling field that is involved in the 36h forecast, at the corresponding time (i.e. for the time $t_1 = t_0 + 24h$, where $t_0, t_1$ are the respective starting dates of the 36h and 12h forecast integrations). A digital filter initialization (DFI) (Lynch and Huang, 1992, and Lynch et al., 1997) is then applied to this coupling field, before the Aladin 12h forecast integration itself.

Compared to the standard NMC method, the implied forecast differences appeared to have some similar variances in the small scales, but much smaller variances in the large scales (see top panels in

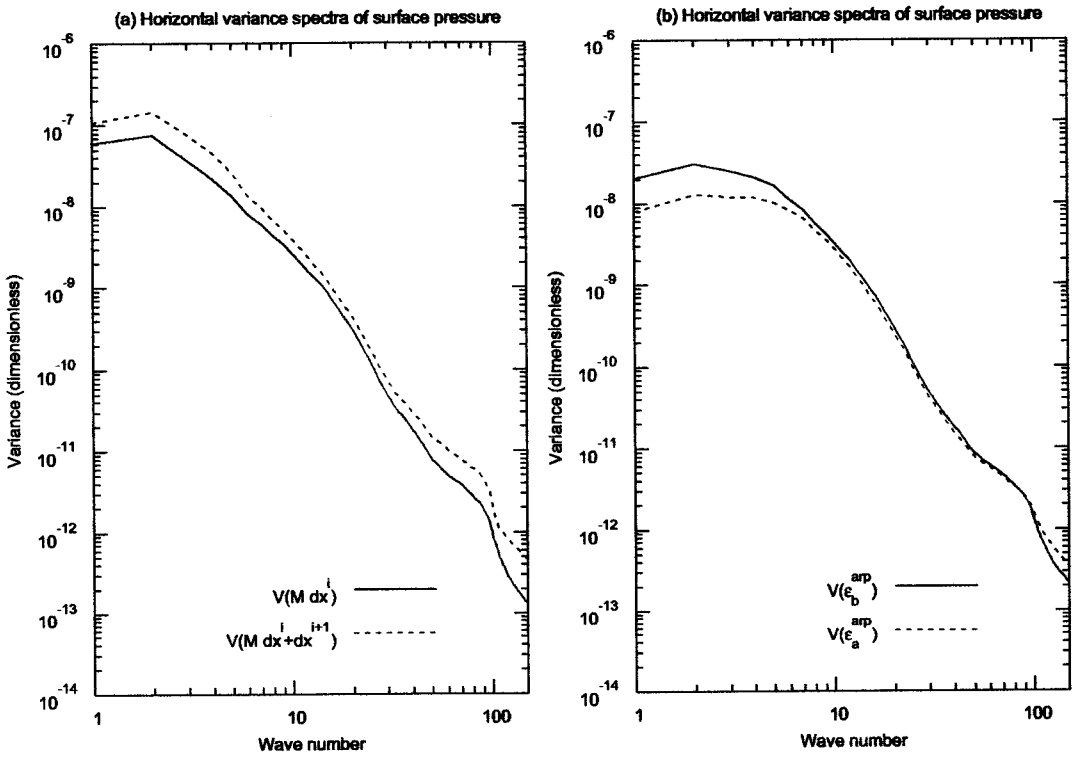


Figure G.4: (a) The evolution of the variance spectra during an analysis step, in the NMC method. This panel shows the variance spectrum of $Mdx^i$ (full line ; this is the first increment, after its six hour evolution), and its modification when adding the second analysis increment $dx^{i+1}$ (dotted line). (b) The evolution of the variance spectra during an analysis step, in the ensemble method. This panel shows the variance spectra of the background differences (full line) and of the analysis differences (dotted line).

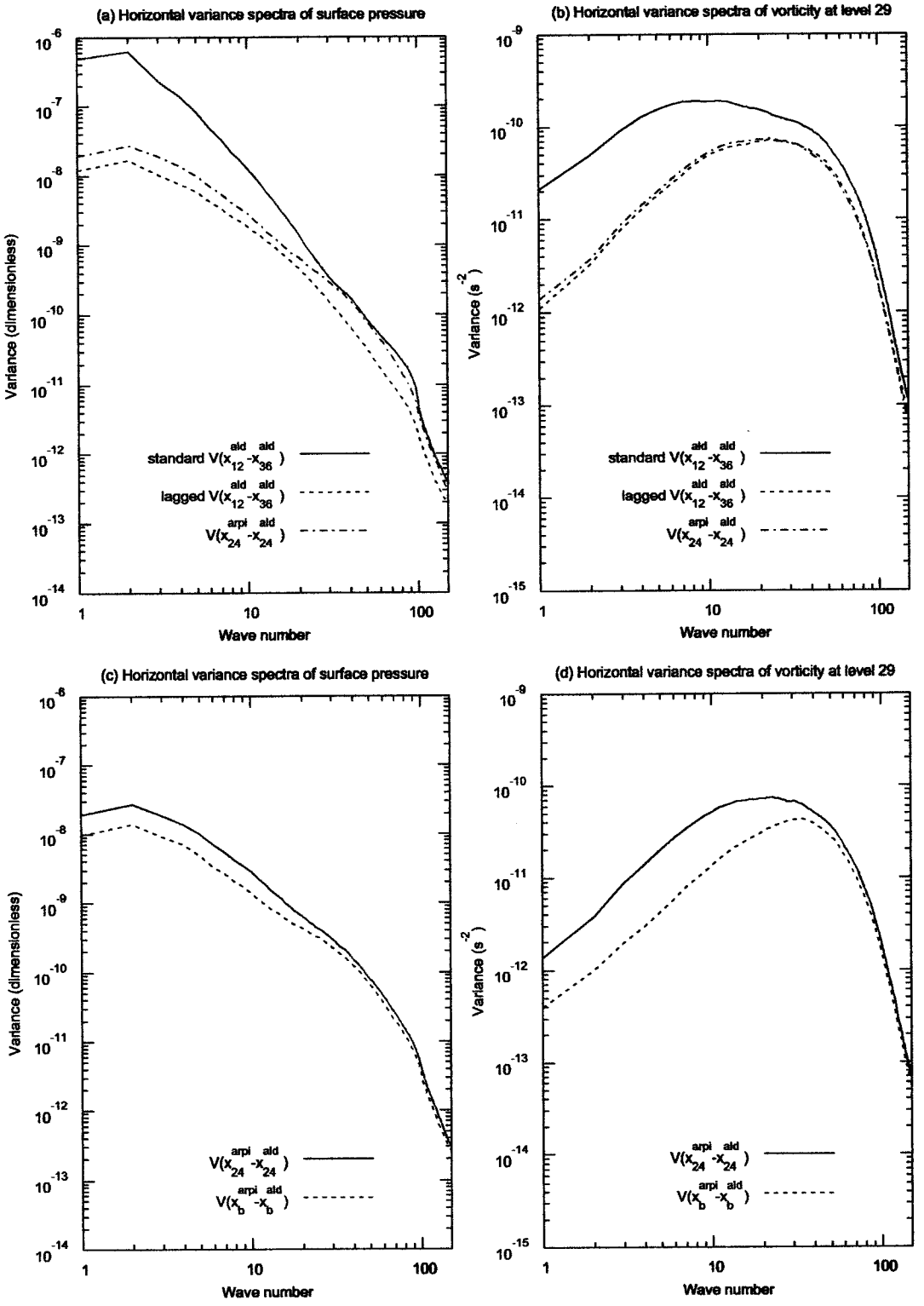


Figure G.5: Top panels: the error variance spectra for the standard (full lines) and lagged (dotted lines) NMC methods, and for the 24 hour Arpège/Aladin differences (dash-dotted lines). Bottom panels: the variance spectra of the 24 hour (full lines) and 6 hour (dotted lines) Arpège/Aladin differences.

figure G.5).

### G.6.2 The link with the Arpège/Aladin model differences

In this section, it will be shown that the lagged NMC method is closely related to the Arpège/Aladin model differences (although this was not explicit in the original description of the lagged NMC method).

The new IC of the 12h forecast $x_{12}^{ald}$ is indeed simply the Arpège 24h forecast $x_{24}^{arp}$, interpolated onto the Aladin grid, and initialized by digital filters:

$$x_{12}^{ald} = \tilde{M}^2 D\tilde{H}\ x_{24}^{arp}$$

where $\tilde{M}$ is the Aladin 6h forecast operator, $D$ is the DFI operator, and $\tilde{H}$ is the interpolation operator of an Arpège field onto the Aladin grid.

By contrast, the 36h forecast $x_{36}^{ald}$ can be seen as another 12h forecast integration, whose IC is an Aladin 24h forecast $x_{24}^{ald}$:

$$x_{36}^{ald} = \tilde{M}^2\ x_{24}^{ald}$$

In other words, the final forecast differences correspond to the 12h forecast evolution of some initial differences, which are the differences between an (initialized) Arpège 24h forecast and an Aladin 24h forecast:

$$x_{12}^{ald} - x_{36}^{ald} = \tilde{M}^2\ (D\tilde{H}\ x_{24}^{arp} - x_{24}^{ald})$$

This means that the initial differences are caused by the differences between the Arpège and Aladin models (which are accumulated over a period of 24 hours):

$$D\tilde{H}\ x_{24}^{arp} - x_{24}^{ald} = (D\tilde{H}M^4 - \tilde{M}^4 D\tilde{H})\ x_a^{arp}$$

where $M$ is the Arpège 6h forecast operator.

Viewing the final forecast differences as being closely related to the Arpège/Aladin model differences may help to interpret some features of the lagged NMC method.

In particular, the correlation spectra of the lagged NMC method are shifted towards the small scales, compared to those of the standard NMC method. This is consistent with the fact that the Arpège/Aladin forecast differences are likely to reflect, to a large extent, the small scale structures that are represented by Aladin, and not by Arpège.

This appears to be confirmed by the examination of the variance spectra of the differences between the (initialized) Arpège 24h forecast and the Aladin 24h forecast (top panels of figure G.5): the large scale variances are much smaller than those of the standard NMC method.

The final 12h forecast integration does not appear to change much the initial variance spectra (except for a small decrease of variance at all scales). This suggests that the essential ingredient of the lagged NMC method is indeed the Arpège/Aladin model differences.

The variance spectra of the Arpège/Aladin 24h differences were also compared to those of the Arpège/Aladin 6h differences (bottom panels of figure G.5). The two additional 18h integrations appear to increase the large scale variances, while the small scale variances remain similar.

### G.6.3 Comparison between the statistics of the three techniques

The variance spectra of the following three error simulation techniques can be compared: the ensemble method, the standard NMC method, and the lagged NMC method (top panels of figure G.6). The spectra of the ensemble method appear to be intermediate between those of the two other estimation techniques.

The smaller amplitude of the ensemble large scale variances, compared to the standard NMC method, is consistent with a more accurate representation (in the ensemble approach) of the influence of the analysis step and of the short forecast ranges (namely six hours).

Moreover, the larger amplitude of the ensemble large scale variances, compared to the lagged NMC method, is consistent with the representation (in the ensemble approach) of the IC and LBC uncertainties. By contrast, the lagged NMC method is in fact based on two assumptions. Firstly, the LBCs are assumed to be perfect. Secondly, the IC uncertainties are assumed to be small: the Arpège and Aladin 24h forecasts are based on the same Arpège analysis, and the small Arpège/Aladin 24h forecast differences may be seen as approximating some rather small IC uncertainties for the additional Aladin 12h forecast.

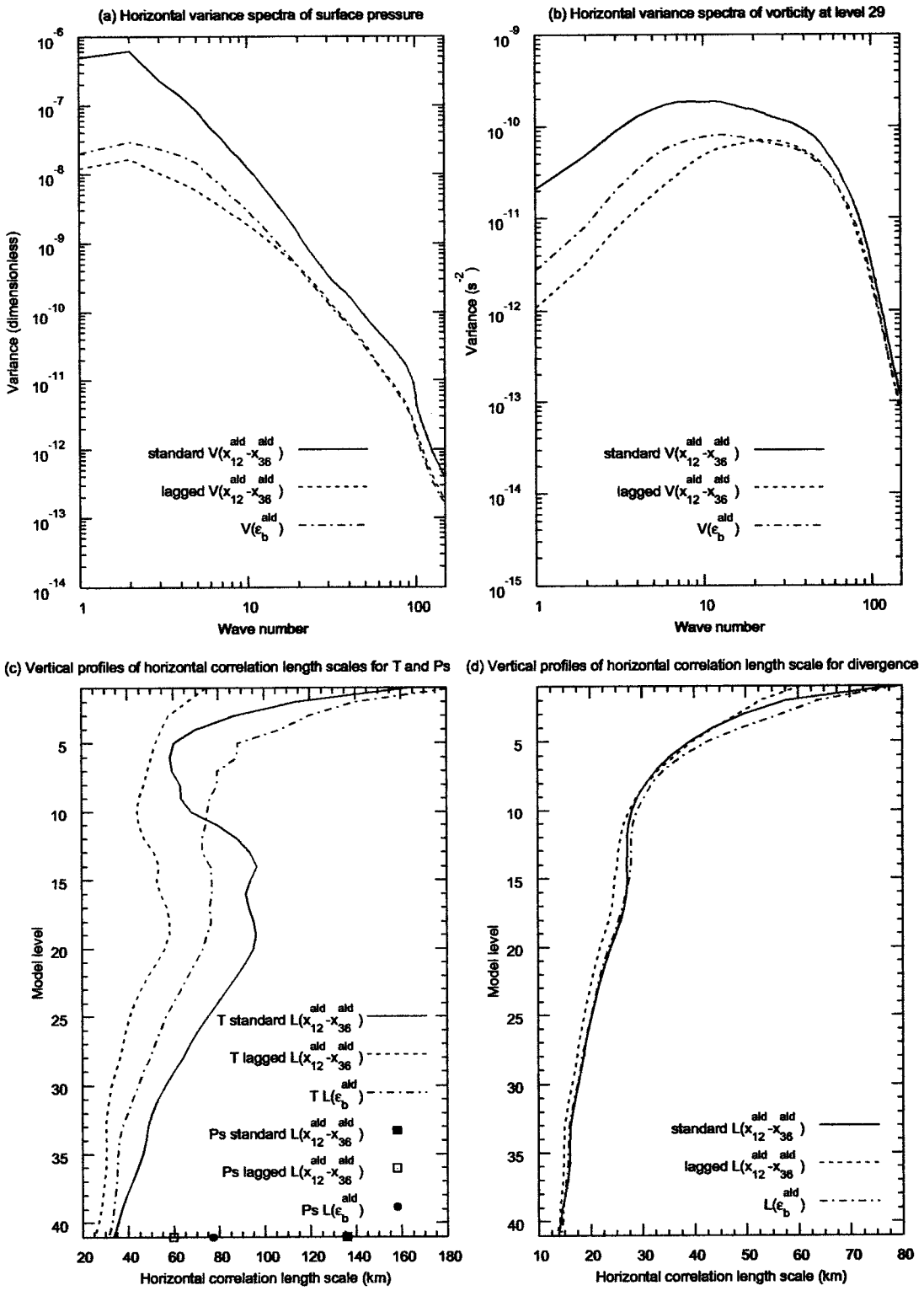


Figure G.6: The error variance spectra (top panels) and the correlation length scales (bottom panels) for the Aladin standard (full lines) and lagged (dotted lines) NMC methods, and for the Aladin background ensemble (dash-dotted lines).

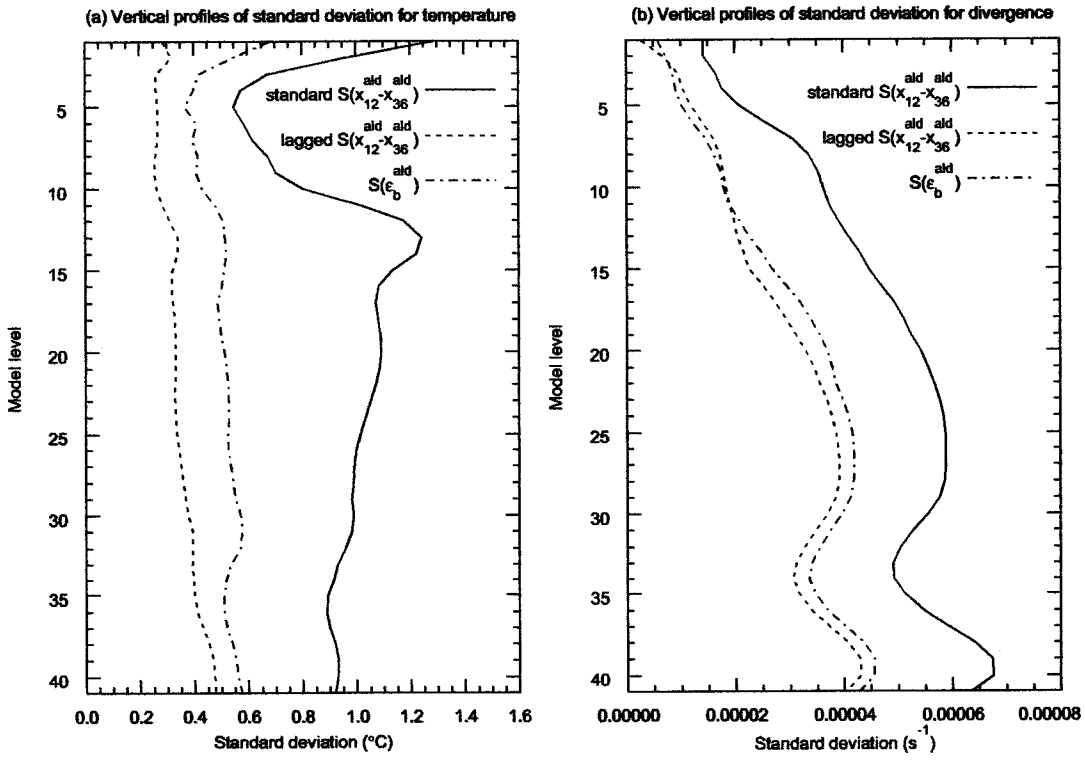


Figure G.7: The vertical profiles of the standard deviations for the Aladin standard (full lines) and lagged (dotted lines) NMC methods, and for the Aladin background ensemble (dash-dotted lines).

The vertical profiles of the length scales (bottom panels of figure G.6), and of the total standard deviations (figure G.7), can be examined also. The results are consistent with the variance spectra: the ensemble results appear to be intermediate between those of the two NMC methods. The contrasts are more pronounced for a large scale variable such as temperature, than for a small scale variable such as divergence.

## G.7 Conclusions and perspectives

An analysis ensemble approach has been compared, formally and experimentally, with two other error simulation techniques. A formal examination of the model state error evolution has been done first, with respect to the influence of the analysis and forecast steps. It appeared in particular that the analysis equation is the very same equation that transforms the background and observation errors into the analysis errors. This is the case, even if the specified covariance matrices are suboptimal, and even if the observation operator is imperfect.

This allows to notice some strong similarities with the evolution of the ensemble differences, in the ensemble technique. It can be shown that the analysis equation is the equation that transforms the background and observation perturbations into the analysis perturbations. This suggests that the influence of the analysis step is well represented in this ensemble technique.

By contrast, the representation of the analysis step can be shown to be inaccurate in the standard NMC method. Indeed, this estimation technique relies essentially on the accumulation and the evolution of some analysis increments. Two results confirm that this process of analysis increment accumulation is not a good approximation of the analysis process.

Firstly, the filtering properties of the gain matrix $\mathbf{K}$ imply that the analysis increment is expected to be of a larger scale than the analysis error. This is supported experimentally by the comparison with the analysis dispersion spectrum. Moreover, this difference appears to be one of the main contributions to the larger scale spectra in the NMC method (compared to the ensemble method).

Secondly, another related experimental result is that the analysis increment accumulation implies an increase of variance for all scales, while the analysis process contributes to a decrease of the large scale dispersion (as expected in particular over a data rich region such as the Aladin-France domain).

The representation of the initial errors and of the analysis effect is even more poor in the so-called lagged NMC method. This variant appeared in fact to be closely related to the Arpège/Aladin model differences. This suggests that some ingredients of the lagged NMC method, such as the latter model differences, may rather be used in order to estimate the contributions of some of the involved model errors. The close link between the lagged NMC method, and the Arpège/Aladin model differences, also explains why the implied forecast difference spectra are shifted towards smaller scales.

The variance spectra of the Aladin background errors provided by the ensemble method were compared to those of the other two techniques. The ensemble spectra were found to be intermediate. Compared to the lagged NMC method, the amplitudes of the large scale variances are increased: this is consistent with the representation, in the ensemble method, of the influence of the initial and lateral boundary conditions uncertainties. These large scale ensemble variances are however smaller than those of the standard NMC method: this is consistent with the more accurate representation (in the ensemble approach) of the analysis step and of the short forecast ranges.

Due to this accurate representation of the analysis and forecast effects, the ensemble method will be used to specify the corresponding error statistics of the Aladin 3D-Var. The latter may include a term which will control the distance to the Arpège analysis (Bouttier, 2002).

Moreover, these Aladin 3D-Var analyses will allow to compare the Aladin background and the Arpège analysis with the available observations. This will also give the possibility to calculate the corresponding departures of the two model states from observations. The statistics of these departures, and more generally, the use of a posteriori diagnostics (e.g. Desroziers and Ivanov, 2001, Sadiki and Fischer, 2005) may also provide some complementary estimates of the error statistics.

The introduction of model perturbations may help to increase further the realism of the ensemble simulation. In a forthcoming publication, a first step towards this will be described, which relies on a detailed examination of the Arpège/Aladin model differences.

Finally, the present study deals with some time- and space-averaged error covariances. Time- and space-dependent features may be explored in the future, e.g. by integrating a larger number of ensemble members, and/or for instance by relating the error structures to the background field structures. Moreover, it may be particularly attractive to use wavelets in this context, both on the globe (Fisher, 2003) for Arpège and in a limited area (Deckmyn and Berre, 2005) for Aladin: this would allow to diagnose and to specify the local error structures.

## G.8 Acknowledgments

The authors would like to thank François Bouttier, François Bouyssel, Claude Fischer, Jean Maziejewski and Alex Deckmyn for carefully reading the early versions of the manuscript. We would like also to thank Robert Mureau for his constructive comments, which have enabled to strengthen the structure of the paper. Simona Ecaterina Ştefănescu and Margarida Belo Pereira are grateful to Météo France for supporting their stays in CNRM/GMAP.

## G.9 Appendix: design of the observation and model perturbations in the analysis ensemble approach

The evolution of the model state errors and of the ensemble differences have been examined in sections G.2 and G.3. The ensemble approach may thus be summarized as follows, e.g. in the perfect model case.

The forecast model is applied to e.g. two perturbed analyses, which provides two backgrounds for the next analysis time. The analysis scheme is then applied to these two perturbed backgrounds, and to two sets of perturbed observations. This provides two new perturbed analyses, and so on. This process ensures that the ensemble perturbations evolve in a similar way as the model state errors.

Beyond this representation of the 6h forecast and analysis effects, some additional choices are introduced, in order to achieve a realistic error covariance estimation.

If one drops the time index notation, the equation of the analysis error can be expressed as follows:

$$e_a = (\mathbf{I} - \mathbf{K}H)e_b + \mathbf{K}e_o$$

where $\mathbf{I}$ is the identity matrix. This implies the following expression for the associated exact error covariance matrices, namely the analysis error covariance matrix $\mathbf{A}_* = \overline{e_a(e_a)^T}$, the background error covariance matrix $\mathbf{B}_* = \overline{e_b(e_b)^T}$ and the observation error covariance matrix $\mathbf{R}_* = \overline{e_o(e_o)^T}$ (if $\overline{e_b(e_o)^T} = 0$, where the overbar corresponds to an ensemble average):

$$\mathbf{A}_* = (\mathbf{I} - \mathbf{K}H)\mathbf{B}_*(\mathbf{I} - \mathbf{K}H)^T + \mathbf{K}\mathbf{R}_*\mathbf{K}^T$$

Similarly, the equation of the analysis ensemble difference (between two members) is the following:

$$\varepsilon_a = (\mathbf{I} - \mathbf{K}H)\varepsilon_b + \mathbf{K}\varepsilon_o$$

which implies the following expression for the associated ensemble difference covariance matrices $\mathbf{A}_\varepsilon = \overline{\varepsilon_a(\varepsilon_a)^T}$, $\mathbf{B}_\varepsilon = \overline{\varepsilon_b(\varepsilon_b)^T}$ and $\mathbf{R}_\varepsilon = \overline{\varepsilon_o(\varepsilon_o)^T}$ (if $\overline{\varepsilon_b(\varepsilon_o)^T} = 0$):

$$\mathbf{A}_\varepsilon = (\mathbf{I} - \mathbf{K}H)\mathbf{B}_\varepsilon(\mathbf{I} - \mathbf{K}H)^T + \mathbf{K}\mathbf{R}_\varepsilon\mathbf{K}^T \tag{G.7}$$

In practice, the two sets of observation perturbations are obtained as random realizations of the Gaussian probability distribution function (pdf) whose mean is zero, and whose covariance matrix corresponds to the specified observation error covariance matrix $\mathbf{R}$:

$$\delta_{o,k}, \delta_{o,l} \sim \mathcal{N}(0, \mathbf{R})$$

i.e. $\overline{\delta_{o,k}(\delta_{o,k})^T} = \overline{\delta_{o,l}(\delta_{o,l})^T} = \mathbf{R}$.

These two random realizations are uncorrelated, which implies that the covariance matrix of the observation difference $\varepsilon_o = \delta_{o,k} - \delta_{o,l}$ is equal to twice the specified covariance matrix $\mathbf{R}$:

$$\begin{aligned} \mathbf{R}_\varepsilon &= \overline{\varepsilon_o(\varepsilon_o)^T} \\ &= \overline{\delta_{o,k}(\delta_{o,k})^T} + \overline{\delta_{o,l}(\delta_{o,l})^T} - \overline{\delta_{o,k}(\delta_{o,l})^T} - \overline{\delta_{o,l}(\delta_{o,k})^T} \\ &= 2\,\mathbf{R} \end{aligned}$$

If the specified observation error covariance matrix $\mathbf{R}$ is a good approximation of the exact observation error covariance matrix $\mathbf{R}_*$, namely $\mathbf{R} \approx \mathbf{R}_*$, then one obtains:

$$\mathbf{R}_\varepsilon = 2\,\mathbf{R}_* \tag{G.8}$$

Similarly, the background difference $\varepsilon_b$ can be seen as the difference between two background perturbations (that are relative to an unperturbed background $x_b$): $\delta_{b,k} = x_{b,k} - x_b$, $\delta_{b,l} = x_{b,l} - x_b$, and $\varepsilon_b = \delta_{b,k} - \delta_{b,l}$.

If these two background perturbations are uncorrelated, and if their spatial covariances are those of the exact background error covariance matrix $\mathbf{B}_*$, then the covariance matrix of $\varepsilon_b$ will be twice the exact error covariance matrix:

$$\mathbf{B}_\varepsilon = \overline{\varepsilon_b(\varepsilon_b)^T} = 2\,\mathbf{B}_* \tag{G.9}$$

By replacing these expressions of $\mathbf{R}_\epsilon$ and $\mathbf{B}_\varepsilon$ in the equation (G.7), one obtains:

$$\mathbf{A}_\varepsilon = 2\,\mathbf{A}_* \tag{G.10}$$

In other words, the covariances of the analysis ensemble differences are expected to be equal to twice the covariances of the analysis errors.

As shown above for the observations, the factor 2 results from the construction of perturbations that are mutually uncorrelated (between different members), and whose (spatial) covariances are made to approximate the error covariances.

A similar principle applies to the possible model perturbations $\varepsilon_m$: the specific model perturbation $\delta_{m,k}$ of a given member $k$ should ideally be mutually uncorrelated with the model perturbation of any other member, and e.g. its variance should approximate the variance of the model error $e_m$.

## References


Bouttier, F., 1994: *Sur la prévision de la qualité des prévisions météorologiques.* PHD dissertation Université Paul Sabatier (Toulouse), 240 pp. [Available from Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex, France].

Bouttier, F. 2002: Strategies for kilometric-scale LAM data assimilation. In: *Report of HIRLAM workshop on variational data assimilation and remote sensing*, Finnish Meteorological Institute, Helsinki, 21-23 January 2002, 9-17. [Available from SMHI, S-60176 Norrköping, Sweden].

Bubnová, R., Horányi, A. and Malardel, S., 1993: International project ARPEGE/ALADIN. *EWGLAM Newsletter* **22**, 117-130.

Daley, R., 1991: *Atmospheric data analysis.* Cambridge Univ. Press, Cambridge, 460 pp.

Deckmyn, A. and Berre, L., 2005: A wavelet approach to representing background error covariances in a limited area model. *Mon. Wea. Rev.* **133**, 1279-1294.

Desroziers, G. and Ivanov, S., 2001: Diagnosis and adaptive tuning of information error parameters in a variational assimilation. *Quart. J. Roy. Meteor. Soc.* **127**, 1433-1452.

Fisher, M. 2003: Background error covariance modelling. In: *Proc. ECMWF Seminar on Recent developments in data assimilation for atmosphere and ocean*, ECMWF, 8-12 September 2003, 45-63. [Available from ECMWF, Reading, RG2 9AX, UK.]

Houtekamer, P. L., Lefaivre, L., Derome, J., Ritchie, H. and Mitchell, H. L., 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.* **124**, 1225-1242.

Lynch, P. and Huang, X.-Y., 1992: Initialization of the HIRLAM Model Using a Digital Filter. *Mon. Wea. Rev.* **120**, 1019-1034.

Lynch, P., Giard, D. and Ivanovici, V., 1997: Improving the Efficiency of a Digital Filtering Scheme for Diabatic Initialization. *Mon. Wea. Rev.* **125**, 1976-1982.

Parrish, D. F. and Derber, J. C., 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.* **120**, 1747-1763.

Radnóti, G., Ajjaji R., Bubnova R., Caian M., Cordoneanu E. and Coauthors, 1995: The spectral limited area model ARPEGE/ALADIN. *PWPR Report Series* **7**, *WMO-TD* **699**, 111-117.

Sadiki, W. and Fischer, C., 2005: A posteriori validation applied to the 3D-Var Arpège and Aladin data assimilation systems. *Tellus* **57A**, 21-34.

Široká, M., Fischer, C., Cassé, V., Brožkova, R. and Geleyn, J.-F., 2003: The definition of mesoscale selective forecast error covariances for a limited area variational analysis. *Meteor. Atmos. Phys.* **82**, 227-244.

Veersé, F. and Thépaut, J.-N. 1998: Multiple-truncation incremental approach for four-dimensional variational data assimilation. *Quart. J. Roy. Meteor. Soc.* **124**, 1889-1908.

Žagar, N., Andersson, E. and Fisher, M. 2004: Balanced tropical data assimilation based on a study of equatorial waves in ECMWF short-range forecast errors. *ECMWF Technical Memorandum* **437**, 30 pp.

# Appendix H

**The evolution of dispersion spectra and the evaluation of model differences in an ensemble estimation of error statistics for a limited area analysis**
***(paper accepted by Monthly Weather Review)***

Simona Ecaterina Ştefănescu[1], L. Berre[2] and M. Belo Pereira[3]
[1]*National Meteorological Administration, SMDCA, Bucharest, Romania*
[2]*Météo-France - CNRM/GMAP, Toulouse, France*
[3]*Instituto de Meteorologia - DOV/CMAR, Lisboa, Portugal*

ABSTRACT

An ensemble of limited area forecasts has been obtained by integrating the Aladin limited area model, in cold-starting mode, from an ensemble of Arpège global analyses and forecasts. This permits error covariances of the Aladin six hour forecast and of the Arpège analysis to be estimated. These two fields may be combined in a future Aladin analysis.

The evolution of dispersion spectra is first studied in a perfect model framework. The Arpège analysis reduces the large scale dispersion of the Arpège background, by extracting some information from observations. Then, the digital filter initialization reduces the small scale dispersion, by removing the noise caused by interpolation of the Arpège analysis onto the Aladin grid. Finally, the Aladin six hour forecast strongly increases the small scale dispersion, in accordance with its ability to represent small scale processes.

Some model error contributions are then studied. The variances of the differences between the Aladin and Arpège forecasts, which are started from the same Arpège analysis, are smaller scale than the Aladin and Arpège perfect model dispersions. The small scale part of these Arpège/Aladin model differences is shown to correspond to structures that are represented by Aladin and not by Arpège. Therefore, this part may be added to the Arpège analysis dispersion. The residual large scale part is more ambiguous, but it may be added e.g. to the Aladin dispersion: this may reflect some effects of the coupling inaccuracies, and strengthen (in a future Aladin analysis) the use of the large scale information from the Arpège analysis.

## H.1 Introduction

In data assimilation, a short range forecast (known as background) is often used as a source of information in addition to observations. The typical range of this forecast is six hours (6h), or even one hour for some regional assimilation schemes. Moreover, the background error covariances are used to filter and propagate the observed information. However, estimating the forecast errors is not trivial, because the true atmospheric state is never exactly known. An ensemble method has been proposed by Houtekamer et al. (1996). This method relies on the time evolution of some perturbations, that are constructed to be consistent with the involved error contributions. There are three basic steps or components that can be involved in the time evolution. First, the analysis scheme is applied to some perturbed observations and to a perturbed background, which provides a perturbed analysis: this simulates the effect of the two information errors (namely the observation errors and the background errors) and of the analysis equation on the initial state uncertainties. Secondly, the perturbed analysis leads to a perturbed forecast, by using the forecast model integration: this aims to reproduce the effects of the atmospheric processes on the evolution of analysis errors into forecast errors. Thirdly, some model perturbations may be added, to reflect the effect of the model errors: this can be done either explicitly at the end of the forecast (or even during the forecast), or implicitly by using different model versions for the different members of the ensemble.

This kind of approach has also been applied at ECMWF by Fisher (2003), using the first two components only, i.e. without adding some model perturbations (partly due to the lack of knowledge about model errors); this version of the method has also been tested at Météo France (Belo Pereira and Berre 2006), to derive statistics for the global model Arpège (Courtier et al. 1991). A similar approach is also presented in Buehner (2005).

While this ensemble method has been used previously for global models, we propose to consider its possible applications in the context of a limited area model. The involved limited area model is Aladin (Bubnová et al. 1993, or Radnóti et al. 1995), which is coupled to the Météo France global model Arpège. The Aladin model is integrated to provide a dynamical adaptation of large scale fields to topography and to other surface characteristics at high resolution. The integration is currently based on a cold-starting mode. This means that its initial condition is basically supplied by the Arpège analysis. An ensemble of Aladin forecasts can thus be produced, with initial conditions and boundary conditions provided by an Arpège ensemble.

There are three main aspects of interest in the experimental results of this ensemble method. The idea is first to study how the ensemble dispersion evolves, through the different steps of the Arpège/Aladin integration. The first step of an Aladin integration is a global Arpège analysis, which is itself a combination of an Arpège 6h forecast with some recent observations. This analysis is then interpolated onto the Aladin grid. This is followed by a digital filter initialization (DFI) (Lynch and Huang 1992) to remove imbalances introduced by the interpolation (Lynch et al. 1997). Finally, a forecast is obtained by integrating the Aladin model, with boundary conditions provided by the Arpège model. In order to understand the statistical features of the Aladin forecast dispersion, it is natural to examine how the dispersion changes through the successive aforementioned processes.

A second point of interest is to study the differences between the Aladin forecasts and the Arpège forecasts that are started from the same Arpège analysis field. This is interesting in itself and also concerning some possible model error representation. Indeed, as mentioned previously, model errors are sometimes simulated by integrating different numerical weather prediction (NWP) models or different versions of a NWP model (e.g., Errico et al. 2001). Thus, it seems interesting to examine the differences between the two NWP models that are involved here, namely the global (low resolution) coupling model (Arpège) and the limited area (high resolution) coupled model (Aladin).

A third issue is the possibility, offered by the ensemble approach, to estimate and compare the error statistics of the Arpège analysis (over the Aladin domain), with the error statistics of

the Aladin forecast. Generally, the global model analysis is thought to be particularly accurate in the large scales, while the limited area forecast tends to be considered to be especially useful in order to provide some small scale information. This general idea is evoked e.g. when trying to design a limited area analysis that accounts both for a limited area model forecast and for an available global analysis (Gustafsson et al. 1997, Brožkova et al. 2001, Bouttier 2002).

The paper is organized as follows. Section H.2 gives a short description of the data set and of notations. In section H.3, the evolution of the dispersion spectra is studied, through the successive basic steps of an Aladin integration. Section H.4 deals with the evaluation and decomposition of the Arpège/Aladin model differences. In section H.5, the implications for the specification of the error statistics of the Aladin forecast and of the Arpège analysis are discussed. Conclusions and perspectives are summarized in section H.6.

## H.2 Description of the data set

The ensemble statistics were computed over a 48 days period, from 4 February 2002 until 23 March 2002. A two-member Aladin ensemble was obtained as follows. First, a two-member ensemble of Arpège assimilation cycles was produced. Each of these two cycles was then used as a source of initial and boundary conditions for the corresponding Aladin experiments. Therefore, the Arpège ensemble is first briefly described, and then the features of the Aladin integrations are presented.

### H.2.1 The Arpège ensemble experiment

The Arpège perturbed analysis and first guess fields were taken from two independent assimilation cycles (Belo Pereira et al. 2002). The basic steps of the Arpège ensemble experiment simulate the time evolution of the Arpège uncertainties, during the data assimilation cycle and in a perfect model framework. Starting from a perturbed analysis that is valid at time $t_0$, a perturbed 6h forecast (valid at time $t_1 = t_0 + 6h$) is obtained by integrating the Arpège model. A perturbed analysis that is valid at time $t_1$ is then obtained by applying the analysis scheme to the perturbed background and to the perturbed observations. The observation perturbations are constructed as random values, which have a Gaussian distribution with a mean equal to zero and a variance equal to the assumed variance of observation errors. The analysis perturbations will thus result from the background and observation perturbations and from the analysis equation. Each perturbed analysis is then integrated in time to provide a perturbed 6h forecast that is valid at time $t_2 = t_1 + 6h$. These basic steps are repeated during successive analysis/forecast cycles.

The very first initial state of the two Arpège experiments is the operational analysis that is valid on 31 January 2002 at 18 UTC: it is therefore the same for the two members, which means that there are no background differences on 1 February 2002 at 00 UTC, at the beginning of the ensemble experiments. The amplitudes of the background differences then grow during the first days, but typically stabilize after three days. Therefore, the period of the data set over which the statistics are calculated starts after this preliminary three days period.

The configuration of the involved Arpège system is based on the operational version of the Arpège model. The corresponding grid is not uniform. It is stretched by a factor $c$, in order to obtain a higher resolution over the European area than over the rest of the globe (Courtier and Geleyn 1988). The stretching factor is equal to $c = 3.5$, and the spectral truncation is T298, which leads to a grid length of about 19 km over the Aladin-France area. There are 41 vertical levels, and the model top is at 1 hPa. The analysis is provided by a four-dimensional variational data assimilation (4D-Var) scheme (Veersé and Thépaut 1998). The assimilated observations are surface pressure from SYNOP, temperature and wind from aircrafts, temperature, wind and humidity from soundings (TEMP radiosondes and PILOT messages), ATOVS AMSU-A

and HIRS cloud-cleared radiances, and atmospheric motion vectors from SATOB. Marginal (in number) data also come from drifting buoys and 10 m wind observations over sea.

### H.2.2 The Aladin integrations and data

In order to obtain an Aladin ensemble, one Aladin 6h integration has been performed in cold-starting mode, from each of the two members of the Arpège ensemble, and for each date of the 48 days period, at 06 UTC and at 12 UTC. The results are similar at 06 UTC and 12 UTC, and they will be shown for fields valid at 12 UTC mostly.

Two additional Aladin members have also been generated (from two other Arpège members), and it has been verified that their time-averaged statistics are quite similar to those of the first two members. This test confirms that the statistics are already stable with the first two members, whose statistical features will be shown in this paper.

The basic steps of an Aladin integration are the following. The Arpège analysis is first interpolated onto the Aladin grid: the Arpège fields are interpolated horizontally (by a twelve points quadratic interpolation; the interpolation points are the twelve nearest points of a 4x4 stencil, which means that the four corner points are not used), and then vertically (in a linear way as a function of pressure) to account for changes in surface pressure (which are induced by changes in the orography). A DFI is applied afterwards. This provides the limited area initial state. Then numerical integrations of the Aladin model are forced by lateral boundary conditions, which are interpolated Arpège forecasts (obtained simply from the Arpège integrations, after the analysis).

These steps were applied to derive a two-member Aladin ensemble from a two-member Arpège ensemble. The interpolation and initialization procedures were also applied to the Arpège 6h forecast fields, for diagnostic purposes. The Aladin model was integrated over the Aladin-France integration domain. The main geometrical characteristics for this integration domain are the following: there are 41 vertical levels as in Arpège (with the model top at 1 hPa), the domain is a square with sides of length $L_x$=$L_y$=2850 km, and the number of grid points in each direction is $J$=$K$=300; the grid resolution is thus $\delta x$=$\delta y$=9.5 km and the spectral truncation corresponds to $M$=$N$=149 ($M, N$ being the maximum wave numbers in each direction). The basic physics and dynamics are the same as in Arpège.

The time-averaged variance spectra of four variables (divergence, vorticity, temperature and surface pressure), at four model levels (13, 22, 29 and 41), were studied for several types of state differences which are listed in section H.2.3. The humidity statistics have not been examined, but they will be the subject of a future study. The ranges which have been studied are 0h and 6h. The model level 13 is located around 225 hPa pressure level, model level 22 corresponds to aproximatively 535 hPa pressure level, model level 29 is located around 775 hPa and model level 41 is located near the surface. The main results will generally be illustrated by the variance spectra of a small scale variable, namely vorticity at level 29, and of a large scale variable, namely the logarithm of surface pressure (in the remainder of the text it will be referred to as surface pressure). In addition, the vertical profiles of standard deviation for temperature and divergence will be studied in section H.3, and also partly in section H.4 for the Arpège/Aladin model differences.

### H.2.3 Notations and terminology

Section H.3 presents results of the Arpège/Aladin assimilation ensemble experiment in a perfect model framework. This allows the contributions of initial condition errors to be simulated under the assumption that models are perfect. Section H.4 deals with a study of 6h accumulated Arpège/Aladin model differences. The corresponding contributions to model errors are discussed further in section H.5. Initial condition and model error contributions are added in section H.5 to provide total error variance estimates, for the Aladin background and also for the Arpège analysis. The corresponding notations will be summarized here.

CONTRIBUTIONS OF INITIAL CONDITION ERRORS (IN A PERFECT MODEL FRAMEWORK)

In section H.3 the time-averaged variances of differences between the two members $k$ and $l$ will be shown at different stages of the Aladin ensemble experiment, such as for the initial state:

$$V(\varepsilon_a^{arpi}) = V(x_a^{arpi}(k) - x_a^{arpi}(l)) \tag{H.1}$$

where for member $k$ for example, the initial state $x_a^{arpi}(k)$ is the Arpège analysis, interpolated onto the Aladin grid, and initialized by DFI:

$$x_a^{arpi}(k) = D\ P\ X_a^{arp}(k) \tag{H.2}$$

$X_a^{arp}(k)$ is the Arpège analysis field over the globe (for member $k$), $P$ is the interpolation operator onto the Aladin grid, and $D$ is DFI. $x_a^{arpi}$ will be refered to as the initialized Arpège analysis. Similarly, the variances of the Aladin 6h forecast differences correspond to:

$$V(\varepsilon_b^{ald}) = V(x_b^{ald}(k) - x_b^{ald}(l)) \text{ with } x_b^{ald}(k) = M^{ald}\ x_a^{arpi}(k) = M^{ald}\ D\ P\ X_a^{arp}(k) \tag{H.3}$$

where $M^{ald}$ is the Aladin 6h forecast operator. The variances of the differences between the Aladin 6h forecasts will be compared in particular to the variances of the differences between the initialized fields of (interpolated) Arpège 6h forecast:

$$V(\varepsilon_b^{arpi}) = V(x_b^{arpi}(k) - x_b^{arpi}(l)) \text{ with } x_b^{arpi}(k) = D\ P\ M^{arp}\ X_a^{arp}(k) \tag{H.4}$$

where $M^{arp}$ is the Arpège 6h forecast operator. $x_b^{arpi}$ will be refered to as the initialized Arpège 6h forecast (or background). Note that the initialization of the Arpège field is indicated by the letter $i$ in the super-script *arpi*. This is introduced to distinguish these fields from the variances of the uninitialized Arpège fields (which will also be studied at the beginning of section H.3), namely:

$$V(\varepsilon_a^{arp}) = V(x_a^{arp}(k) - x_a^{arp}(l)) \text{ with } x_a^{arp}(k) = P\ X_a^{arp}(k) \tag{H.5}$$

$$V(\varepsilon_b^{arp}) = V(x_b^{arp}(k) - x_b^{arp}(l)) \text{ with } x_b^{arp}(k) = P\ M^{arp}\ X_a^{arp}(k) \tag{H.6}$$

CONTRIBUTIONS OF MODEL ERRORS

In section H.4, the time-averaged variances of the differences between the Aladin and (initialized) Arpège 6h forecasts (noted by $\varepsilon_m^{ald-arpi}$), which are started from the same Arpège analysis field (i.e., differences between the Aladin and Arpège forecasts which belong to the same member), will be studied:

$$V(\varepsilon_m^{ald-arpi}) = V(x_b^{ald} - x_b^{arpi}) = \frac{1}{2}[\ V(x_b^{ald}(k) - x_b^{arpi}(k)) + V(x_b^{ald}(l) - x_b^{arpi}(l))\ ]$$
$$\text{with } x_b^{ald}(k) - x_b^{arpi}(k) = (M^{ald}DP - DPM^{arp})X_a^{arp}(k) \tag{H.7}$$

For simplicity, all operators have been assumed to be linear. The model difference expression $M^{ald}DP - DPM^{arp}$ indicates that $\varepsilon_m^{ald-arpi}$ corresponds to 6h accumulated differences between the Aladin and Arpège models. The examination of such model differences is similar to the approach in Errico et al. (2001) for instance.

Cohn and Dee (1988, page 592, equation (2.6c)) and Dee (1995, page 1131, equation (5)) also discuss the estimation of model error by using model differences. While a projection operator of a low-resolution to a high-resolution representation is used in our study, these two references present a similar approach in which a projection in the opposite direction (i.e., a truncation

operator transforms a high-(or infinite-) resolution representation to a lower resolution representation) is performed.

In practice, the respective model differences statistics for members $k$ and $l$ are very similar to each other as expected (since they describe the basic differences between the Arpège and Aladin integrations, independently from the details of their initial conditions). They were averaged to formally preserve an equal statistical weight for both members in the examined statistics.

$V(\varepsilon_m^{ald-arpi})$ will also be compared with $V(x_b^{ald})$ and with $V(x_b^{arpi})$, which are variances of the full forecast fields. As for $V(\varepsilon_m^{ald-arpi})$, they are calculated as an average of the respective results for members $k$ and $l$. The detailed expression of e.g. $V(x_b^{ald})$ is thus the following:

$$V(x_b^{ald}) = \frac{1}{2}[\, V(x_b^{ald}(k)) + V(x_b^{ald}(l)) \,] \tag{H.8}$$

$V(x_b^{ald})$ and $V(x_b^{arpi})$ describe the atmospheric variability, as seen by the Aladin and (initialized) Arpège forecast fields respectively.

In section H.5.1, the respective model error variances of Arpège and Aladin will be estimated from $V(\varepsilon_m^{ald-arpi})$. They will be denoted $V(\varepsilon_m^{arpi})$ and $V(\varepsilon_m^{ald})$ respectively.

TOTAL ERROR VARIANCE ESTIMATES

A total error variance $V(\varepsilon_{bm}^{ald})$ will be defined as follows for the Aladin 6h forecast:

$$V(\varepsilon_{bm}^{ald}) = V(\varepsilon_b^{ald}) + V(\varepsilon_m^{ald}) \tag{H.9}$$

This amounts to summing up the following two contributions to the Aladin 6h forecast error variance: the errors induced by initial condition uncertainties (under the assumption that the models are perfect), and the Aladin model errors. This is analogous to the Kalman filter equations such as described in Daley (1991, equations (13.3.5) p.376 and (13.3.7) p.377).

Similarly, a total error variance $V(\varepsilon_{am}^{arpi})$ for the (initialized) Arpège analysis will be defined as follows, to account for the two contributions to the corresponding uncertainties:

$$V(\varepsilon_{am}^{arpi}) = V(\varepsilon_a^{arpi}) + V(\varepsilon_m^{arpi}) \tag{H.10}$$

## H.3 Contributions to the evolution of dispersion spectra

The dispersion of the Aladin 6h forecasts results from a rather complex evolution: the contributions to this evolution are studied in the current section, by examining the spectral and vertical variations of the dispersion, after the successive steps of the Aladin integration.

### H.3.1 The effect of the Arpège analysis

The effect of the Arpège analysis can be diagnosed by comparing the respective dispersions of the Arpège analysis and of the Arpège background (both being first projected onto the Aladin grid). The corresponding horizontal variance spectra for surface pressure and vorticity at level 29 and the vertical profiles of standard deviations for temperature and divergence are presented in figure H.1.

The analysis dispersion appears to be smaller than the background dispersion. This is consistent with the expected reduction of uncertainty, when assimilating observations to correct the Arpège background.

This reduction of dispersion is somewhat stronger in the large scales than in the small scales, which is particularly visible for surface pressure. This is consistent with the large amplitude of background errors in the large scales (compared to observation errors) for the main observed variables, such as surface pressure, temperature, radiances and wind. The analysis is expected to correct more efficiently background components whose error amplitude (compared to the

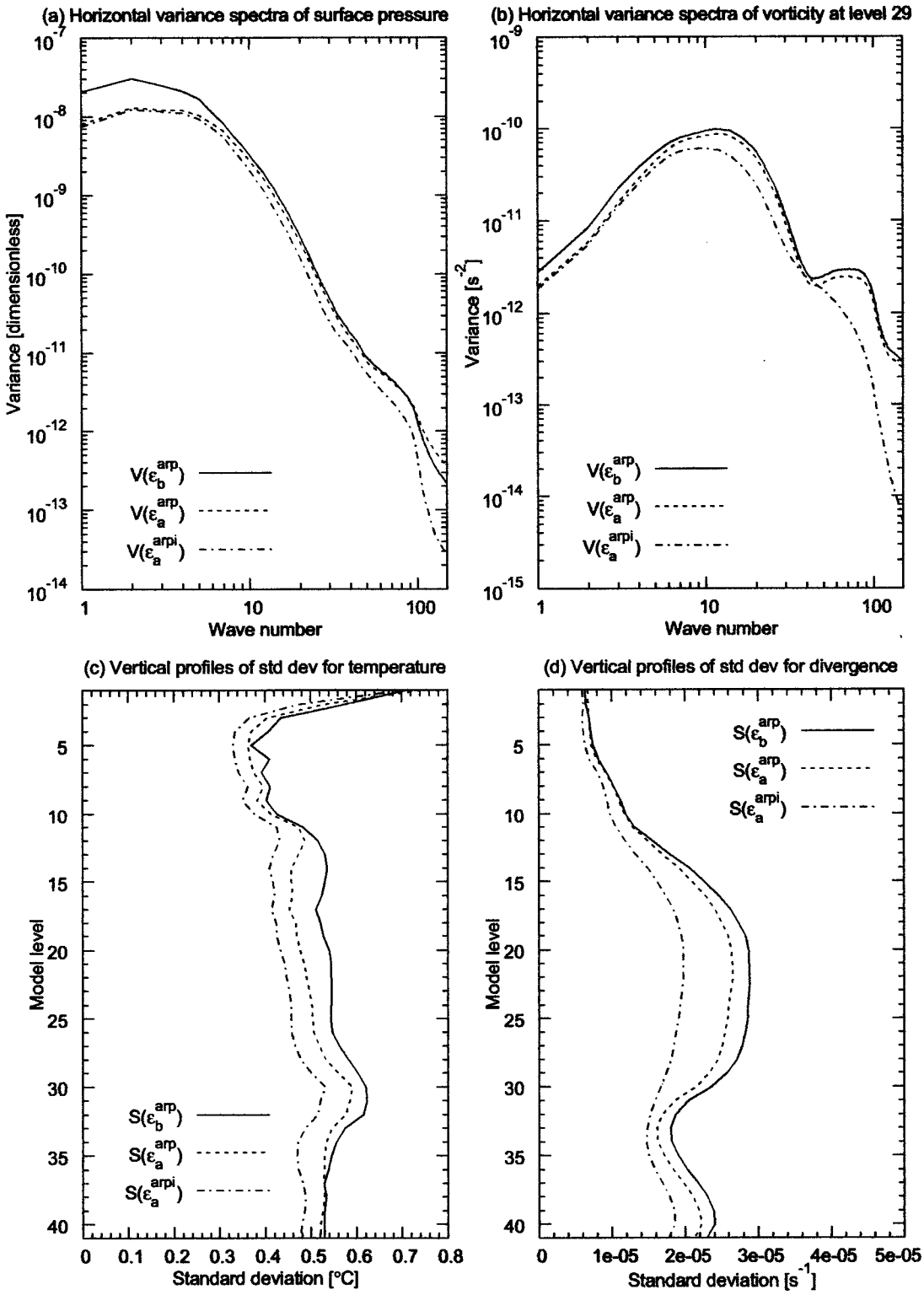


Figure H.1: The respective effects of the Arpège analysis (seen on the Arpège uninitialized fields) and of the DFI: the solid, dotted, and dash-dotted lines correspond to the dispersion statistics of the Arpège uninitialized background, uninitialized analysis, and initialized analysis respectively.

corresponding observation error amplitude) is large, as discussed also in Daley (1991, the final paragraph of section 4.5 on page 130 and figure 5.9 on page 174) and Daley and Ménard (1993, figure 2 on page 1558).

Similarly, the reduction of dispersion is also more pronounced in the mid-troposphere than near the surface, which may be linked with the greater importance of large scale features at these middle levels. The analysis effect is smaller at the highest levels, probably due to the lower observation density there.

The resolution of the Arpège model over the Aladin-France integration domain is about 19 km, which corresponds roughly to wave number 40. The shape of the dispersion spectra for wave numbers larger than 40 is therefore completely determined by the interpolation itself (there is no information about these small scales from the Arpège model itself). The peak of the small scale variance at wave numbers 70-90, which is particularly visible for vorticity in figure H.1, is therefore likely to correspond to some artificial noise that is introduced by the interpolation. Such interpolation effects are consistent with those observed in previous studies relying on the NMC method (Sadiki et al. 2000). As discussed in this latter reference, they are related to the fact that the interpolation is essentially mathematical (rather than physical), compared for instance to what a high resolution NWP model can provide. In particular, physically, the only meaningful information in the interpolation is the assumption of hydrostatic balance in the vertical.

### H.3.2 The effect of the DFI

The effect of the DFI (applied on the Aladin grid) can be examined by comparing the respective dispersions of the (uninitialized) Arpège analysis and of the initialized Arpège analysis (figure H.1). The dispersion of the initialized fields is smaller than for the uninitialized fields. The comparison of the dispersion spectra indicates moreover that the reduction of dispersion is rather small in the large scales, and that it increases towards the small scales.

It can be observed in particular that the artificial peak of the variance for vorticity in the small scales, which was present in the uninitialized analysis, has been removed by the DFI. The DFI thus appears to remove the noise that is introduced by the interpolation. The reduction of the dispersion is particularly large for divergence in the mid-troposphere. At these levels, the artificial peak of variance was strong.

For diagnostic purposes, it is interesting to apply DFI to the Arpège 6h forecast field as well: it allows to visualize the Arpège analysis effect, without the contribution of the interpolation noise itself (figure H.2). Note that the Arpège 6h forecast is the field used as background for the Arpège analysis, and therefore the Arpège analyses and forecasts are valid at the same time. A comparison of figures H.1 and H.2 indicates that the analysis effect on the initialized background is similar to the analysis effect on the uninitialized background. The dispersion variance is indeed reduced by the analysis, as expected.

### H.3.3 The effect of the Aladin 6h forecast

Similarly, in order to visualize the effects of the atmospheric evolution during the Aladin 6h forecast, it is convenient to compare the dispersion statistics of the Aladin 6h forecast with those of the Arpège initialized fields. It is better to use the Arpège initialized fields (instead of the Arpège uninitialized ones), as DFI is able to remove the main part of the interpolation noise.

A first way to evaluate the effect of the 6h integration of Aladin is thus to compare the respective dispersion statistics of the Aladin initial state (which is an Arpège analysis initialized by DFI) and of the Aladin 6h forecast (figure H.3). Note that the latter field is valid 6h later than the former. A general increase of the variance can be observed. The large scale increase of variance likely reflects the increase of the dispersion induced by some atmospheric instabilities

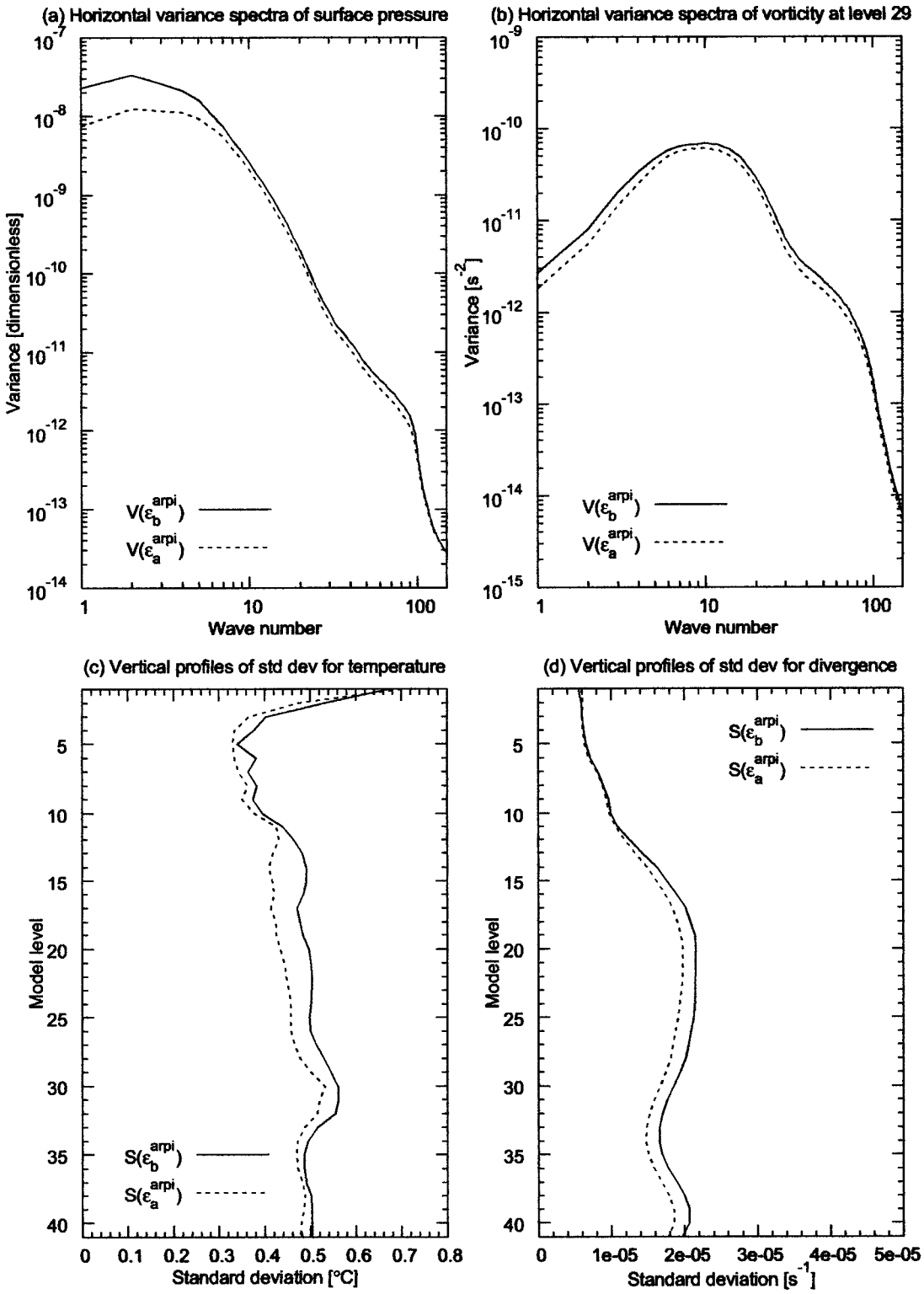


Figure H.2: The effect of the Arpège analysis (seen on initialized fields): the solid and dotted lines correspond to the dispersion statistics of the Arpège (initialized) background and (initialized) analysis respectively.

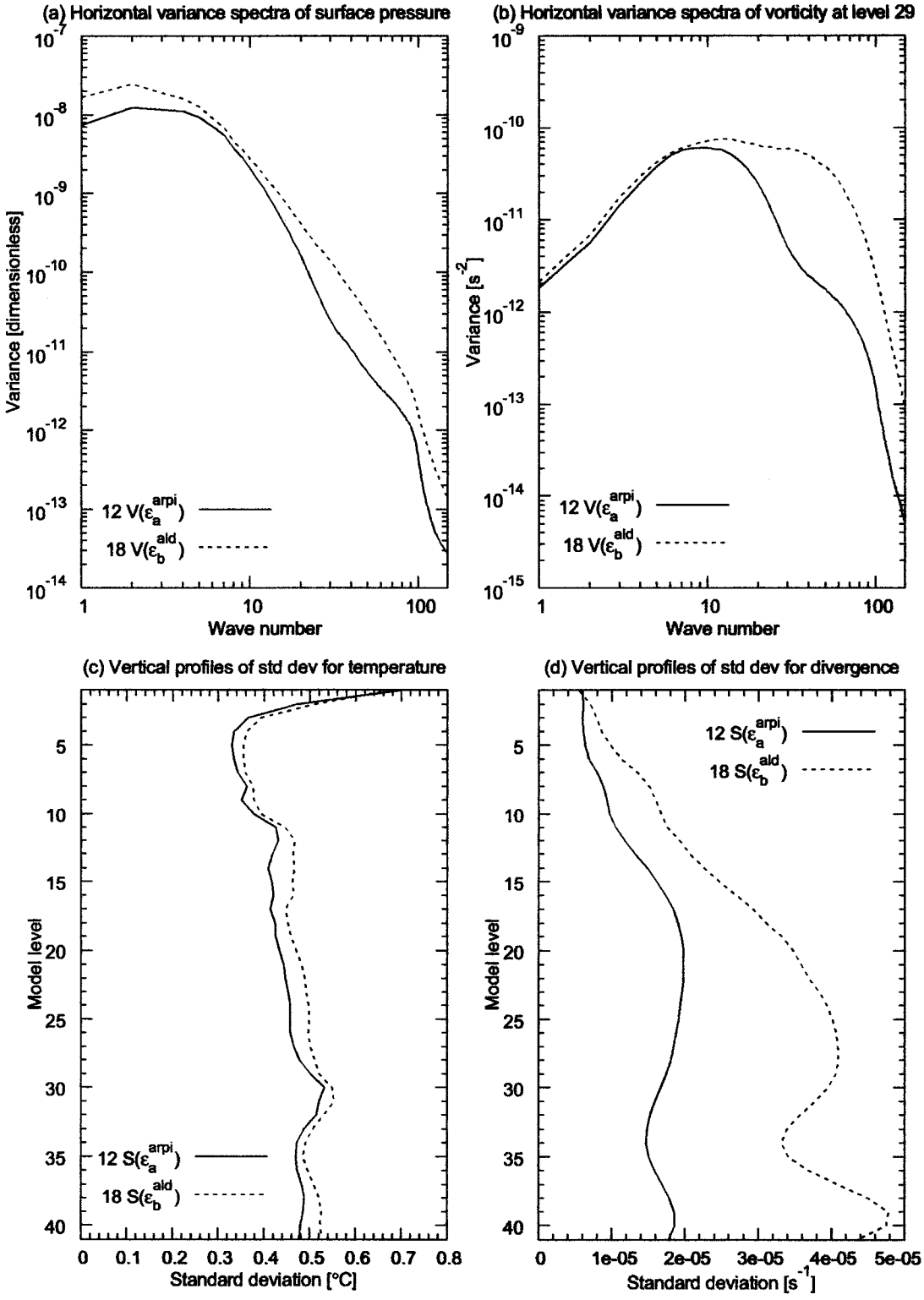


Figure H.3: The effect of the Aladin 6h forecast: this figure shows the evolution of the Arpège initialized analysis dispersion (valid at 12 UTC, solid lines) into the Aladin 6h forecast dispersion (valid at 18 UTC, dotted lines).

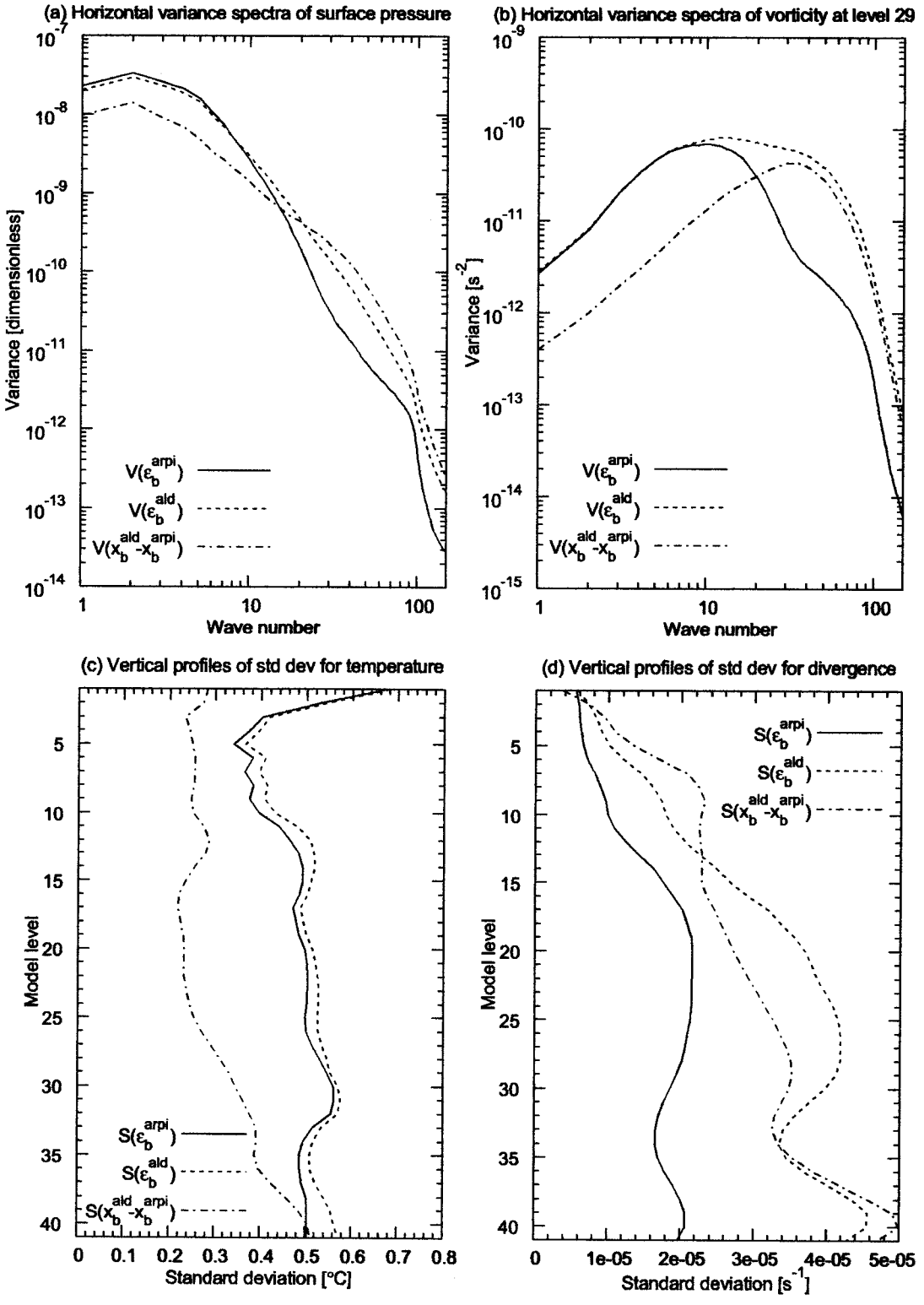


Figure H.4: The perfect model dispersions of the Arpège initialized 6h forecast (solid lines) and of the Aladin 6h forecast (dotted lines); the dash-dotted lines represent the statistics of the differences between the Aladin 6h forecasts and the Arpège (initialized) 6h forecasts, which correspond to the same ensemble member.

(such as baroclinic developments for instance). Moreover, probably the increase of the small scale variance is induced by the higher resolution of the Aladin model.

To confirm this, it is useful to compare also the respective dispersion statistics of the Arpège (initialized) 6h forecast and of the Aladin 6h forecast (figure H.4). It has been noticed previously that the DFI step implies essentially the removal of some small scale interpolation noise. In contrast, as can be seen especially from the vorticity variance spectrum, the Aladin 6h forecast provides a relatively large amount of additional small scale energy. This is consistent with the representation of small scale dynamical effects in the Aladin model, as discussed in Sadiki et al. (2000). The increase of small scale energy is somewhat more pronounced at low levels, which may be related to the higher resolution topography, and more generally to the larger importance of small scale processes at low levels.

The increase of the small scale dispersion by the Aladin 6h forecast implies an increase of the total variance. As expected, this effect is much stronger for a small scale variable such as divergence than for a large scale variable such as temperature (see bottom panels of figures H.3 and H.4).

The smaller amplitude of the small scale variance in the Arpège 6h (initialized) forecast field is expected to some extent, due to the lower resolution of the Arpège model. The latter implies naturally that the contributions of some small scale processes, that exist on the Aladin grid, are dissipated by the Arpège numerical diffusion, or not even represented explicitly in Arpège. On the other hand, this result may also be seen as unrealistic, knowing that the dispersion is supposed to reflect the involved uncertainties: one could consider in fact that the smaller variances of the Arpège forecasts suggest that the Arpège forecast is more accurate than the Aladin forecast in the small scales, which would be rather unexpected knowing that the Arpège model has a lower resolution. In order to study these issues more deeply, differences between the Arpège and Aladin forecasts will be examined in the next section.

## H.4 Model differences evaluation and decomposition

In this section, differences between Aladin and (initialized) Arpège 6h forecasts, which are started from the same Arpège analysis field, will be studied. As mentioned in section H.2.3, this amounts to examining the influence of the 6h accumulated model differences between Arpège and Aladin. First, the nature of the involved model differences will be summarized. Secondly, the resulting forecast differences will be examined and compared to the Arpège and Aladin perfect model dispersions (which were studied in section H.3). Thirdly, the respective variance spectra of the Arpège and Aladin full forecast fields and their difference will be studied. This will allow to highlight the strong contribution of the associated forecast amplitude differences in the small scales. Finally, this will lead to consider a scale decomposition of the variance of the Arpège/Aladin forecast differences.

### H.4.1 The nature of the Arpège/Aladin model differences

Three basic features are expected to contribute to the differences between Arpège and Aladin forecast fields on the Aladin grid, when they are started from the same Arpège analysis.

Firstly, the difference of horizontal resolution means that some small scale structures are either not represented at all by Arpège, or dissipated by Arpège numerical diffusion.

Secondly, Arpège is a global model, while Aladin is a limited area model. The Aladin coupling is based on a Davies-Kâlberg relaxation scheme (Davies 1976), with a three-hour coupling frequency. Any coupling technique is imperfect, which means that the large scale information which is provided by Arpège may be distorted.

Thirdly, Arpège fields are interpolated onto the Aladin grid, in order to obtain fields which can be used by Aladin. As evidenced in section H.3.1, this interpolation is not perfect, and likely produces some unrealistic small scale features beyond the Arpège resolution.

The interpolation noise was also shown to be reduced by DFI in section H.3.2. Moreover, in the remainder of the paper, we will concentrate on initialized Arpège fields. This means that the horizontal resolution and the coupling are the main two factors contributing to the Arpège/Aladin forecast differences.

The basic physics and dynamics are the same in both models, consequently the Arpège/Aladin differences are not expected to provide a full model error simulation. On the other hand, this allows us to focus on the aforementioned specific differences and on their implications for the estimation of error statistics. In other words, this is a first natural step in order to go beyond the perfect model approach.

### H.4.2 Comparison of model differences with the Arpège and Aladin perfect model dispersions

Figure H.4 presents the spectral and vertical distribution of the Arpège/Aladin forecast differences, together with the respective perfect model dispersions of the Arpège and Aladin 6h forecasts (which were studied in section H.3). The examination of the vorticity variance spectrum indicates that the dispersion of the Arpège/Aladin differences is more shifted towards the small scales than the Arpège and Aladin dispersions, with a maximum around wave numbers 30-40 (instead of 10-20 for the two other curves). This is expected, given that the Arpège and Aladin models are based on similar equations and that the model differences result to a large extent from the resolution differences. The vertical profiles presented in figure H.4 (bottom panels) show larger standard deviation values of the Arpège/Aladin differences at low levels (than in the mid- and high-troposphere): this is consistent with the effect of the higher resolution topography, and more generally with the larger importance of small scale processes at low levels.

On the other hand, it can be noticed that the non-zero variance values of the Arpège/Aladin differences do not concern only the scales that are resolved by Aladin and not by Arpège. The Aladin-specific scales correspond indeed to wave numbers that are larger than 40. This is visible in particular in the variance spectra of surface pressure: the variance of the Arpège/Aladin differences appears to be maximum in the largest scales, rather than in the small or intermediate scales. It may therefore be interesting to identify more specifically the part of the Arpège/Aladin differences that corresponds to resolution differences. This is the object of the next section.

### H.4.3 Decomposition of the Arpège/Aladin differences

The difference of resolution implies that some small scale structures will be represented by Aladin and not by Arpège (either not at all, or dissipated by the Arpège numerical diffusion).

In order to estimate the contribution of these small scale structures, one may first express the variance of the Arpège/Aladin difference field, as a function of the variances of the two full forecast fields:

$$V(\varepsilon_m^{ald-arpi}) = V(x_b^{ald} - x_b^{arpi}) = V(x_b^{ald}) + V(x_b^{arpi}) - 2\ cov(x_b^{ald}, x_b^{arpi}) \qquad \text{(H.11)}$$

where $V(x_b) = \overline{x_b x_b^T}$ is the variance of the full forecast field $x_b$, and $cov(x_b^{ald}, x_b^{arpi}) = \overline{x_b^{ald}(x_b^{arpi})^T}$ is the covariance between the two forecast fields. As mentioned in connection with equation (H.8), note that $V(x_b)$ describes the full atmospheric variability, as seen by the forecast field $x_b$.

It is possible to rewrite equation (H.11) as:

$$V(x_b^{ald} - x_b^{arpi}) = V(x_b^{ald})\ (1 + \beta - 2\sqrt{\beta}\rho) \qquad \text{(H.12)}$$

where $\beta = \frac{V(x_b^{arpi})}{V(x_b^{ald})}$ is the ratio between the Arpège variance and the Aladin variance (which can also be seen as the percentage of the Aladin variance that is represented by Arpège), and $\rho$ is the correlation between the Aladin and Arpège forecast fields. The parameter $\beta$ describes

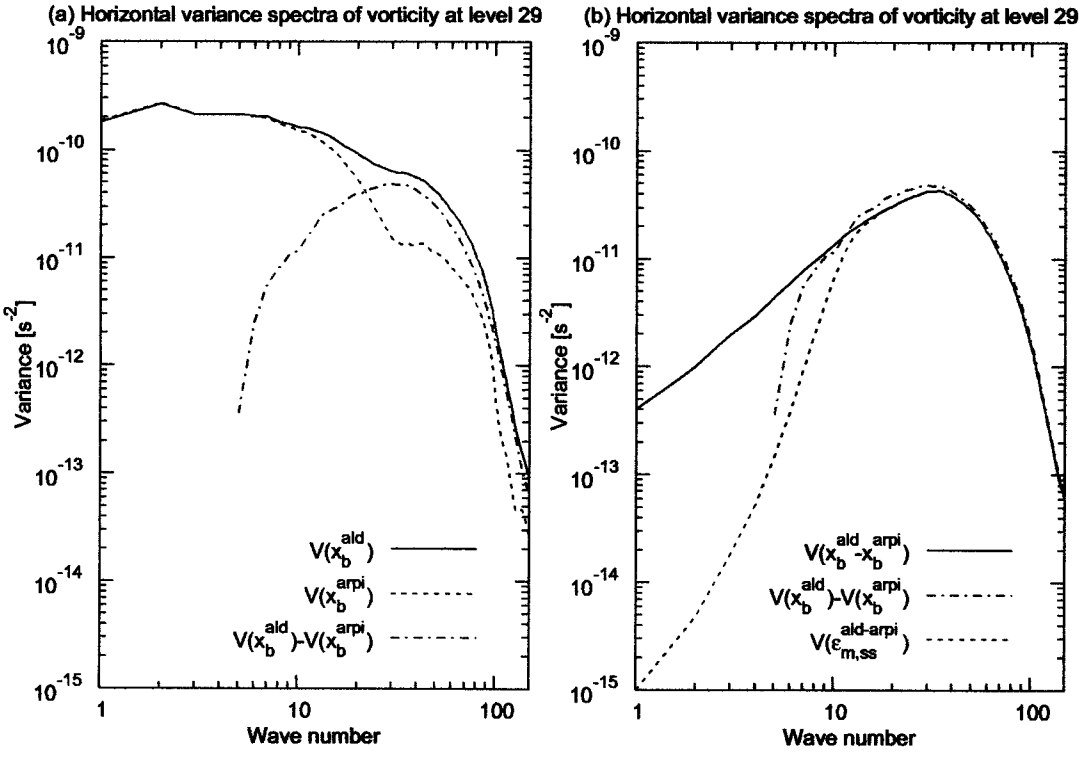


Figure H.5: The difference between the Arpège and Aladin variances (dash-dotted lines). Left panel: comparison with the variances of the Arpège (dotted line) and Aladin (solid line) forecast fields. Right panel: comparison with the variance of the Arpège/Aladin differences (solid line). Missing values in the largest scale range of the curve $V(x_b^{ald}) - V(x_b^{arpi})$ correspond to negative values. In the right panel, the small scale part of the variance of the Arpège/Aladin differences is also shown (dotted line).

the influence of the average amplitude differences, while the parameter $\rho$ corresponds to the influence of the average phase differences. For instance, $\beta$ is likely to be close to zero for small scale structures which are dissipated by Arpège, and $\rho$ is likely to be close to zero if the Arpège and Aladin waves are in quadrature. These two contributions may be partly distinguished, by rearranging the equation (H.12) as the sum of the following two terms:

$$V(x_b^{ald} - x_b^{arpi}) = V(x_b^{ald})\ (1-\beta) + V(x_b^{ald})\ 2\beta\ (1 - \frac{\rho}{\sqrt{\beta}}) \qquad \text{(H.13)}$$

where the first term of the right hand side can be easily shown to correspond to the difference between the full forecast variances of Aladin and of Arpège, namely $V(x_b^{ald})\ (1-\beta) = V(x_b^{ald}) - V(x_b^{arpi})$.

This term has been calculated and is plotted in figure H.5, where it can be compared with the variances of the Arpège and Aladin full forecast fields $V(x_b^{arpi})$ and $V(x_b^{ald})$ (left panel). The Arpège and Aladin variances are very similar in the large scales, while the Aladin variance is much larger than the Arpège variance in the small scales, as expected (see the beginning of the last paragraph in section H.3.3). This implies that the difference between these two variance spectra will correspond to a spectrum that has its maximum in the small scales.

This difference between the Arpège and Aladin variances can also be compared with the variance of the Arpège/Aladin differences (right panel of figure H.5). The two curves look very similar in the small scales (for wave numbers larger than 10), which means:

$$V(\varepsilon_m^{ald-arpi}) \sim V(x_b^{ald}) - V(x_b^{arpi}) \qquad \text{(H.14)}$$

In other words, the (energy) amount of the Arpège/Aladin forecast differences tends to be the

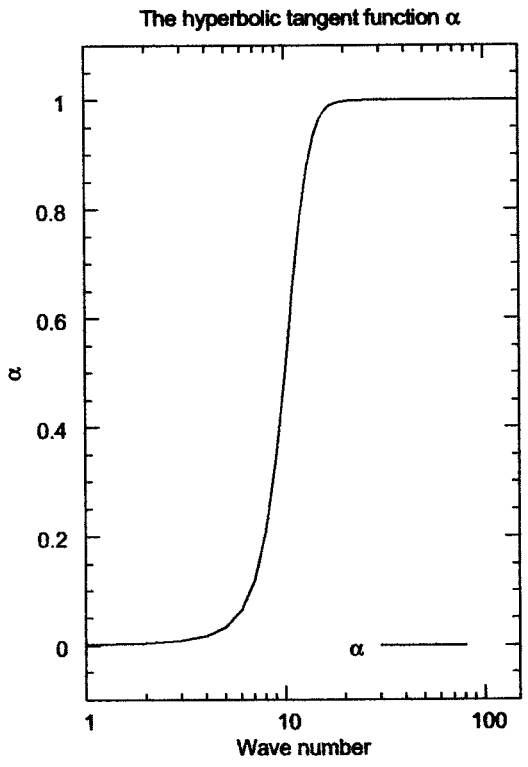


Figure H.6: The hyperbolic tangent function $\alpha(k^*) = 0.5\ (\ tanh[(k^* - 10)/3]\ +\ 1)$.

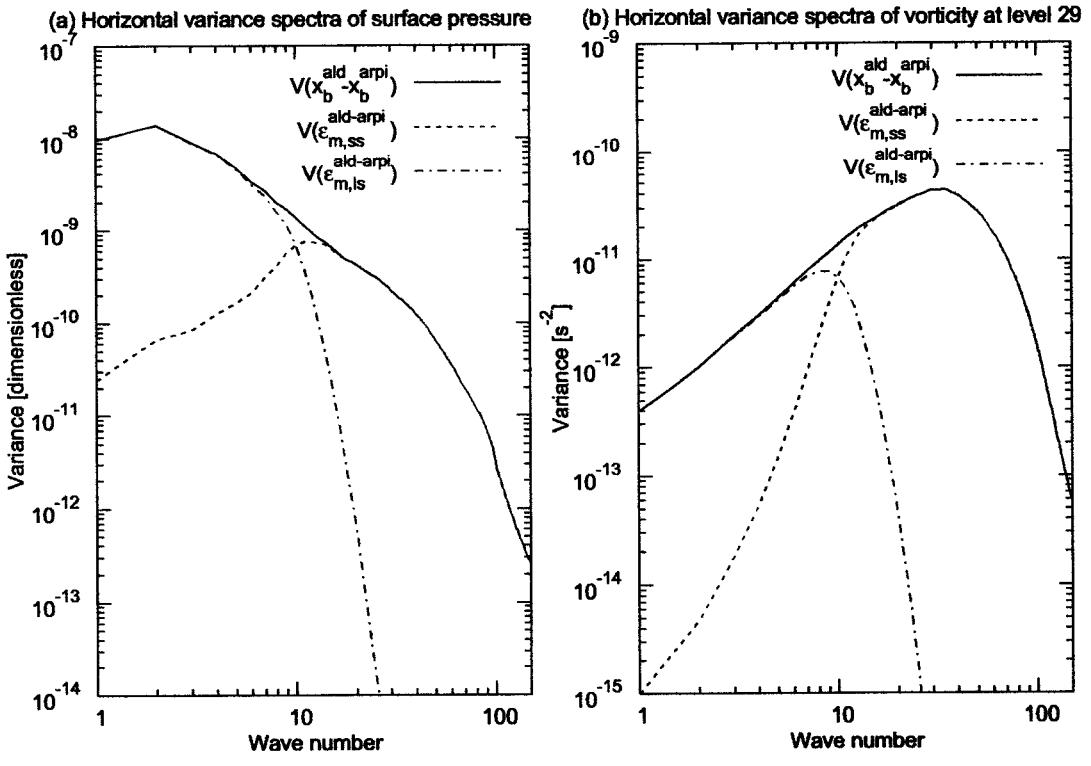


Figure H.7: The decomposition of the variance of the Arpège/Aladin differences (solid lines), as a function of a small scale part (dotted lines) which is induced by resolution differences, and of a residual large scale part (dash-dotted lines).

same as the (energy) amount of Aladin structures which are not represented by Arpège. This is consistent with the expectation that the small scale structures which are represented by Aladin and missed by Arpège predominate in the Arpège/Aladin forecast differences. Therefore, one may conclude that the small scale Arpège/Aladin forecast differences correspond essentially to some resolution-induced Arpège model errors with respect to the Aladin grid.

It should also be mentioned that, in general, some possible small scale noise of the interpolation (of Arpège fields onto the Aladin grid) is another potential contribution to the variance of the Arpège/Aladin forecast differences. Nevertheless, there are two results which indicate that this is not a major contribution in the present case, which deals with initialized Arpège fields. Firstly, while some spurious noise is visible in the small scale spectrum of uninitialized Arpège fields, the DFI appears to remove most of it (top panels of figure H.1). Secondly, if the possible remaining interpolation noise had been a major small scale contribution to the Arpège field $x_b^{arpi}$ and to the associated Arpège/Aladin forecast differences $\varepsilon_m^{ald-arpi}$, one would expect a strong statistical link between the structures of $x_b^{arpi}$ and those of $\varepsilon_m^{ald-arpi}$. This can be measured by the amplitude of the corresponding cross-covariance $cov(x_b^{arpi}, \varepsilon_m^{ald-arpi})$ (which can be compared to $V(\varepsilon_m^{ald-arpi})$ and $cov(x_b^{ald}, \varepsilon_m^{ald-arpi})$ for instance). In fact (see appendix A), the result $V(\varepsilon_m^{ald-arpi}) \sim V(x_b^{ald}) - V(x_b^{arpi})$ (right panel of figure H.5) indicates that $cov(x_b^{arpi}, \varepsilon_m^{ald-arpi}) \sim 0 \ll cov(x_b^{ald}, \varepsilon_m^{ald-arpi}) \sim V(\varepsilon_m^{ald-arpi})$. Therefore, the small scale structures of $\varepsilon_m^{ald-arpi}$ appear to be much more linked to the structures of $x_b^{ald}$. This is consistent with the idea that $\varepsilon_m^{ald-arpi}$ corresponds essentially to some (resolution-induced) structures which are represented by Aladin and which are missed by Arpège.

The sharper decrease of $V(x_b^{ald}) - V(x_b^{arpi})$ towards the large scales (compared to the corresponding decrease of $V(\varepsilon_m^{ald-arpi})$) suggests moreover that there may be some large scale components in the Arpège/Aladin differences that do not correspond to the former resolution-induced structures and differences in the variance amplitudes. Therefore, one may decompose the variance of the Arpège/Aladin differences as the sum of two components:

$$\underbrace{V(x_b^{ald} - x_b^{arpi})}_{V(\varepsilon_m^{ald-arpi})} = \underbrace{\alpha\, V(x_b^{ald} - x_b^{arpi})}_{V(\varepsilon_{m,ss}^{ald-arpi})} + \underbrace{(1-\alpha)\, V(x_b^{ald} - x_b^{arpi})}_{V(\varepsilon_{m,ls}^{ald-arpi})} \qquad \text{(H.15)}$$

where $\alpha$ is a parameter that varies from 0 in the large scales to 1 in the small scales. $V(\varepsilon_{m,ss}^{ald-arpi})$ corresponds thus to a small scale part of the Arpège/Aladin differences, which are induced by the resolution differences. $V(\varepsilon_{m,ls}^{ald-arpi}) = V(x_b^{ald} - x_b^{arpi}) - V(\varepsilon_{m,ss}^{ald-arpi})$ is a residual, that corresponds rather to a large scale part of the Arpège/Aladin differences.

The parameter $\alpha$ may be modelled by a simple hyperbolic tangent function, as a function of the two-dimensional (2D) horizontal wave number $k^*$: $\alpha(k^*) = a\ (\ tanh[(k^* - b)/c] + d)$, with $a = 0.5$, $b = 10$, $c = 3$ and $d = 1$ (figure H.6). These values were chosen in order to have variations between 0 and 1, and in order to ensure that $V(\varepsilon_{m,ss}^{ald-arpi}) = \alpha V(x_b^{ald} - x_b^{arpi})$ has a similar shape as $V(x_b^{ald}) - V(x_b^{arpi})$ (right panel of figure H.5).

The results of the decomposition are shown in figure H.7. It illustrates how the two parts contribute differently to the variance of the Arpège/Aladin differences as a function of scale.

## H.5 Implications for the specification of the Arpège and Aladin error statistics

### H.5.1 Implied changes in the variance spectra

The possibility of estimating some model error contributions $V(\varepsilon_m^{arpi})$ and $V(\varepsilon_m^{ald})$ will now be considered, on the basis of the variance of the Arpège/Aladin differences and of its decomposition (as discussed in section H.4). The resulting total error variances $V(\varepsilon_{am}^{arpi})$ and $V(\varepsilon_{bm}^{ald})$ of the

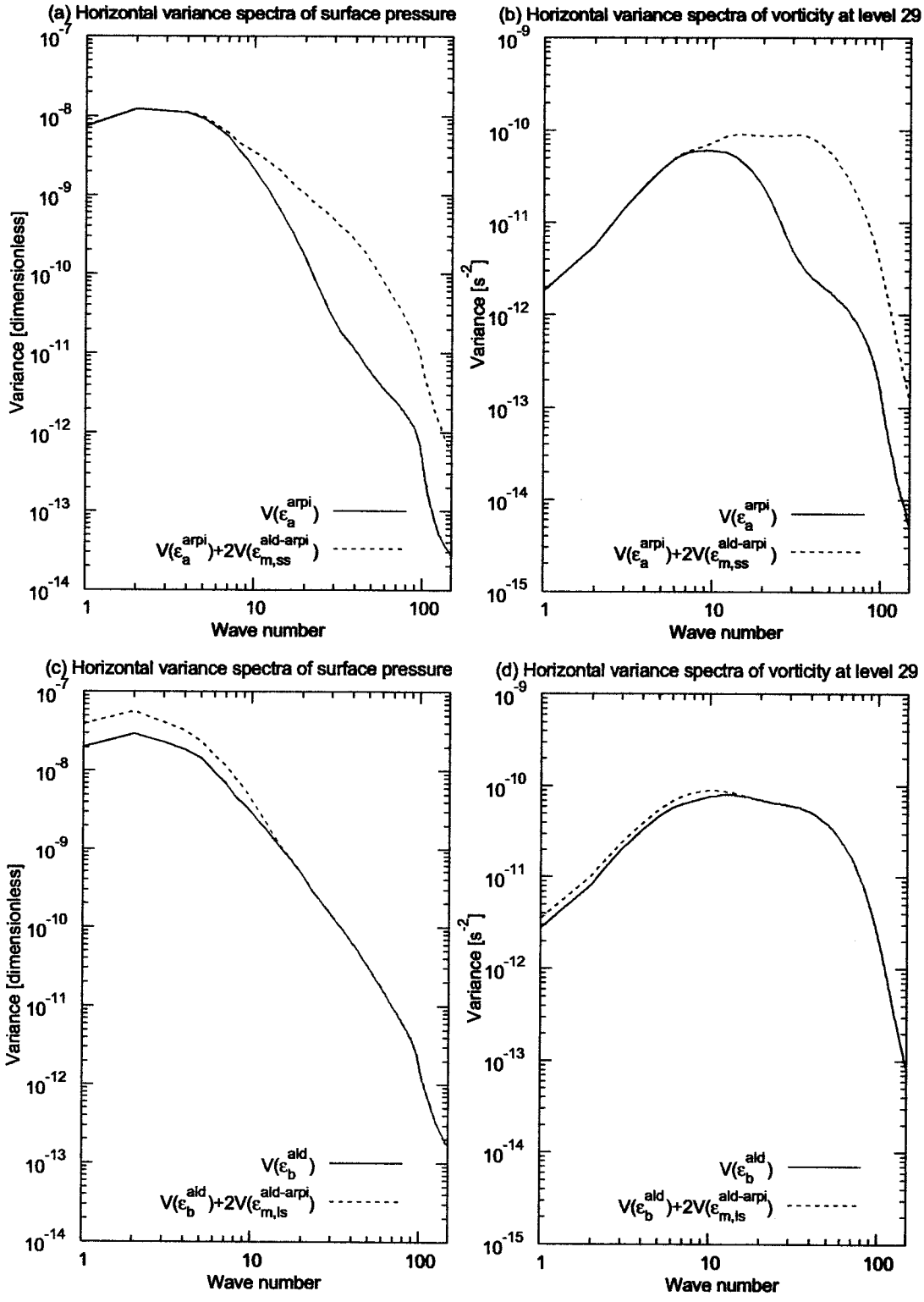


Figure H.8: Top panels: the variance spectra of the Arpège analysis, before (solid lines) and after (dotted lines) adding the contribution of the small scale part of the model differences. Bottom panels: the variance spectra of the Aladin 6h forecast, before (solid lines) and after (dotted lines) adding the contribution of the large scale part of the model differences.

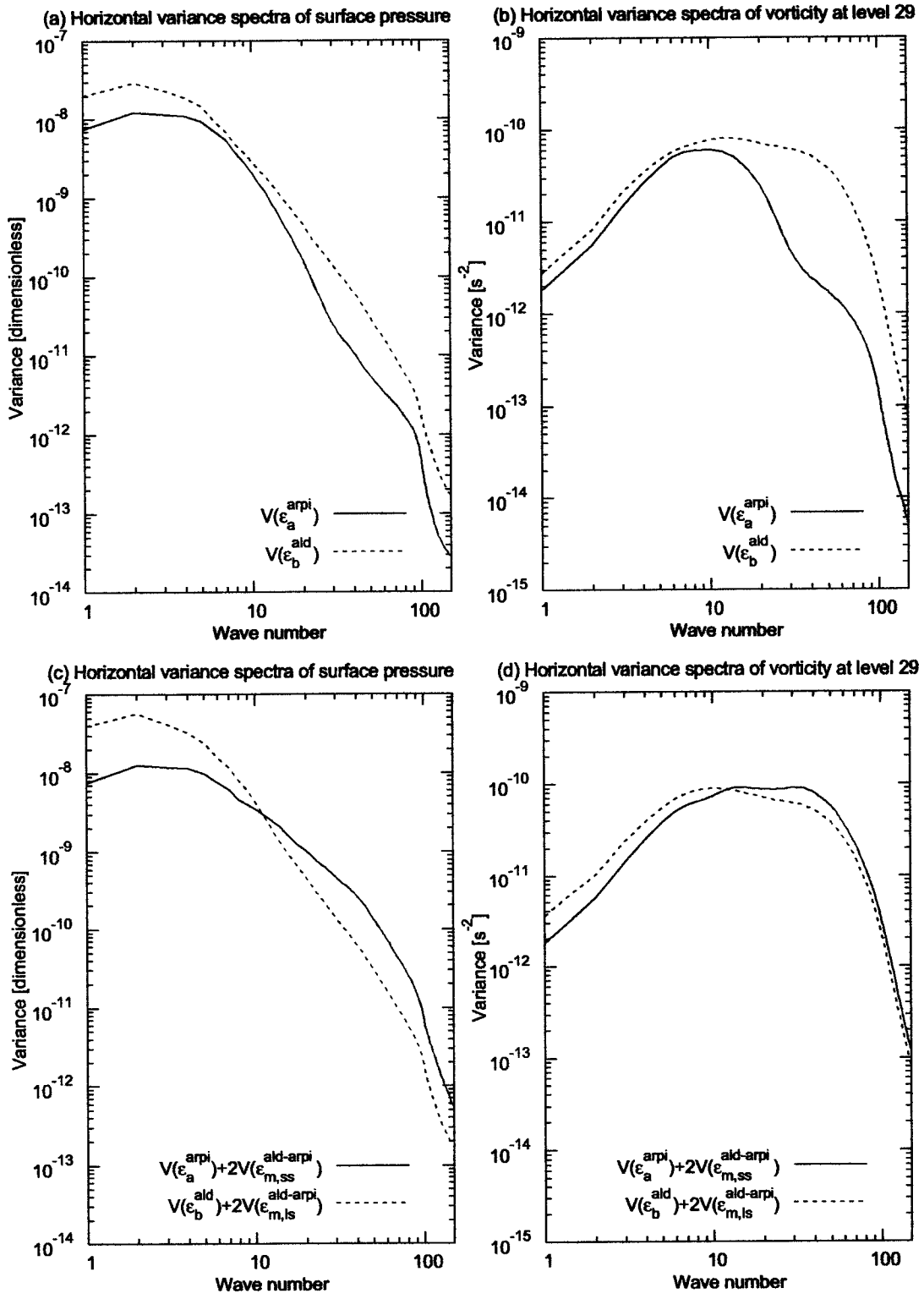


Figure H.9: Top panels: the variance spectra of the Arpège analysis (solid lines) and of the Aladin 6h forecast (dotted lines), before the addition of the contributions of the model differences. Bottom panels: the "total" variance spectra of the Arpège analysis (solid lines) and of the Aladin 6h forecast (dotted lines) (i.e., after adding the model differences contributions, according to the first scenario).

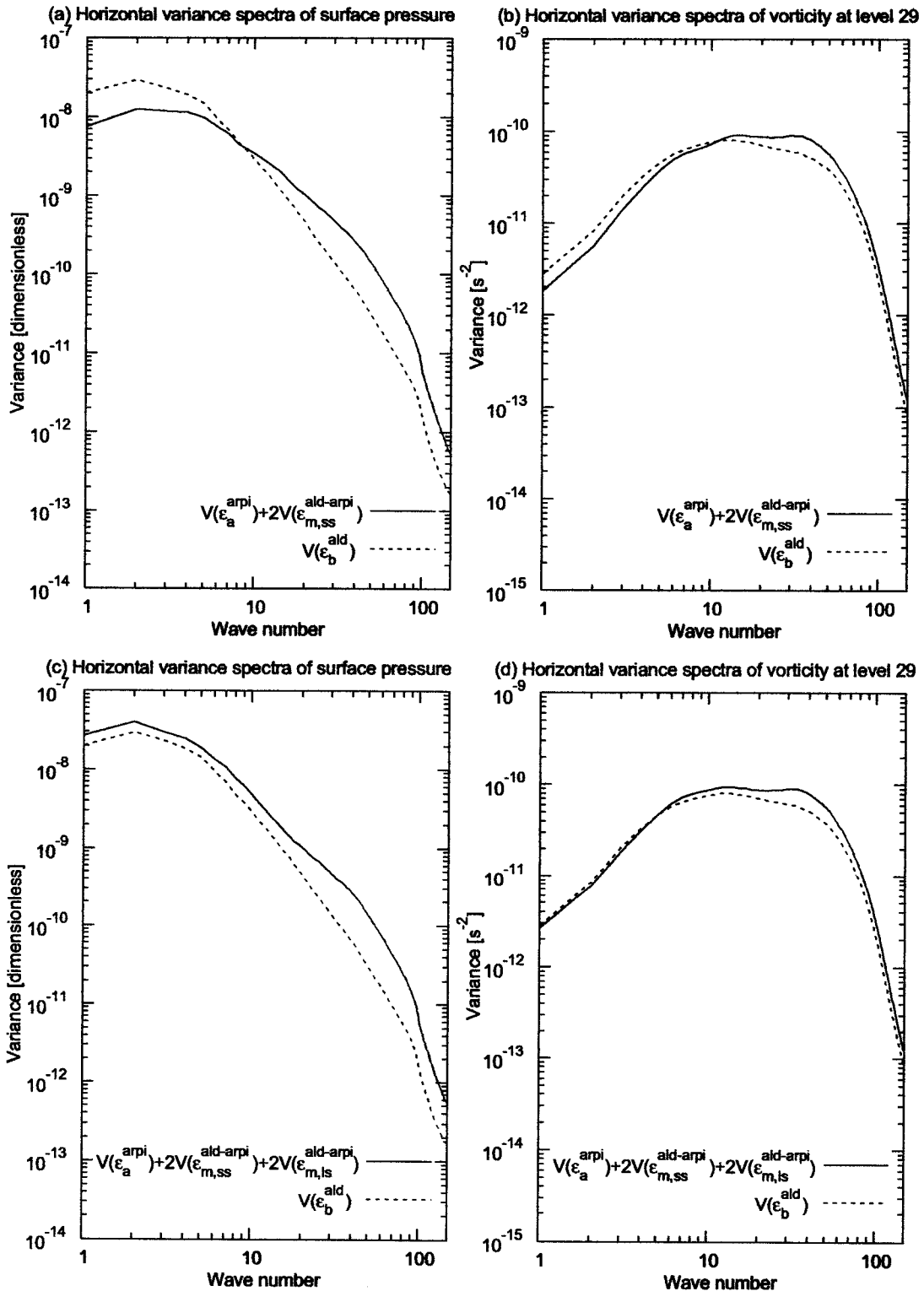


Figure H.10: The "total" variance spectra of the Arpège analysis (solid lines) and of the Aladin 6h forecast (dotted lines), according to the other two scenarios. Top panels: when the contribution of the large scale part of the model differences is added neither to the Aladin dispersion, nor to the Arpège dispersion. Bottom panels: when the contribution of the large scale part of the model differences is added to the Arpège dispersion.

Arpège (initialized) analysis and of the Aladin 6h forecast respectively will also be examined, following equations (H.9) and (H.10) of section H.2.3. The former field is the actual initial state of the Aladin forecast, when Aladin is integrated in cold-starting mode; the latter field is a possible background for an Aladin three-dimensional variational data assimilation (3D-Var). These two fields may also be combined, either within a variational analysis, or through a blending technique (Brožkova et al. 2001).

The small scale part of the model differences variance, $V(\varepsilon_{m,ss}^{ald-arpi})$, appeared to be closely related to the small scale structures that are represented by Aladin, and which are missed by Arpège. Therefore, it can be considered that these small scale variances correspond to some Arpège model errors, with respect to the truth on the Aladin grid:

$$V(\varepsilon_{m}^{arpi}) = 2V(\varepsilon_{m,ss}^{ald-arpi}) \tag{H.16}$$

It may be mentioned that the contribution of the small scale model differences, namely $2V(\varepsilon_{m,ss}^{ald-arpi})$, can be considered to be applicable equally to the Arpège 6h forecast and to the Arpège analysis: this is due to the same low resolution of both Arpège fields. Another remark is that the factor 2 in equation (H.16) yields a proper weight for this model error contribution, compared to the contribution of the dispersion arising from the perturbations of the Arpège assimilation cycle. The detailed equations are presented in the appendix B, and the explanation can be described briefly as follows. The assimilation dispersion corresponds to the variance of the difference between two perturbed fields, involving two independent realizations of the perturbations: the resulting variance can easily be shown to be twice the variance of the basic perturbations. In contrast, the small scale part of an Arpège/Aladin model difference field may be seen as a single realization of the associated model error probability distribution function (pdf): a factor 2 is therefore to be used, in order to recover a contribution that is twice the variance of the involved model error.

The implied changes in the variance spectra are shown in the top panels of figure H.8 (where $V(\varepsilon_{a}^{arpi})$ and $V(\varepsilon_{am}^{arpi})$ can be compared). While the large scale variance is practically unchanged, the small scale variance is strongly increased, as expected.

The large scale part of the model differences variance, $V(\varepsilon_{m,ls}^{ald-arpi})$, is more difficult to interprete. It is likely to correspond to a mixture of several effects. One may consider in particular two contributions to these large scale model differences.

A first contribution may come from the interactions between the small scales and the large scales. This corresponds to the non linear effects of resolving smaller scale features on the simulation of the large scale solution. One may consider, in particular, that a more realistic simulation of small scale processes may also contribute to improve the simulation of the large scale phenomena, if the involved processes are non linear.

A second contribution may come from the coupling inaccuracies. Any coupling technique is imperfect, which means that the large scale information that is provided by the Arpège model may be distorted.

By themselves, the Arpège/Aladin model differences, and the proposed scale decomposition, are insufficient to distinguish the two aforementioned possible contributions. In order to go further, one may consider some additional experiments in the future. For instance, one could integrate some Aladin forecasts that have the same (low) resolution as the Arpège forecasts. Statistics of the differences between the high and low resolution Aladin forecasts may thus be calculated: this would provide an estimate of the contribution of the Arpège/Aladin resolution differences to the Arpège/Aladin forecast differences variance. Statistics of the differences between the low resolution Aladin forecasts and the Arpège forecasts may also be calculated: this would provide an estimate of the contribution of the coupling inaccuracies to the Arpège/Aladin forecast differences variance.

Due to these two possible contributions, three different scenarios will be considered in the remainder of this paper. A first possible scenario is to consider that the large scale model differ-

ences provide an (upper bound) estimate of the contribution of the Aladin coupling inaccuracies:

$$V(\varepsilon_m^{ald}) = 2V(\varepsilon_{m,ls}^{ald-arpi}) \tag{H.17}$$

The implied changes of the Aladin variance spectra are shown in the bottom panels of figure H.8 (where $V(\varepsilon_b^{ald})$ and $V(\varepsilon_{bm}^{ald})$ can be compared). As expected, the small scale variance is essentially the same, while the large scale variance is significantly increased.

The implications of this first scenario are illustrated in figure H.9. This figure represents the comparison between the variance spectra of the Arpège analysis and of the Aladin forecast. When the contributions of the model differences are not considered at all (top panels), the Aladin variances appear to be larger than the Arpège variances, for all considered scales. In the large scales, this is consistent with the reduction of error due to the Arpège analysis. But it does not look realistic in the small scales, when considering that the dispersion is supposed to describe the involved uncertainties, and knowing that Aladin has higher resolution than Arpège.

When the contributions of the model differences are added to the Arpège and Aladin dispersions, as described above, the situation changes. The small scale variances of Arpège are now larger than those of Aladin: this is consistent with the lower Arpège resolution. Moreover, the representation of the larger uncertainty of the Aladin forecast in the large scales has been strengthened.

The second scenario is similar to the first one, except that the variance of the large scale model differences is no longer added to the Aladin dispersion (or to the Arpège dispersion): $V(\varepsilon_m^{ald}) = 0$ and therefore $V(\varepsilon_{bm}^{ald}) = V(\varepsilon_b^{ald})$. Compared to the previous scenario, this implies essentially that the uncertainty of the Aladin forecast in the large scales is reduced (top panels of figure H.10), although it remains larger than the Arpège analysis uncertainty.

A third scenario is to consider that the large scale model differences variance (*in addition to the variance of the small scale ones*) should be added to the Arpège dispersion (by considering e.g. that they correspond rather to the beneficial effects of the Aladin higher resolution on the Aladin large scale simulation): $V(\varepsilon_m^{arpi}) = 2V(\varepsilon_{m,ls}^{ald-arpi}) + 2V(\varepsilon_{m,ss}^{ald-arpi})$. Compared to the other two scenarios, this would imply for vorticity (bottom right panel of figure H.10) that the Arpège variance in the large scales would be much closer to (although still smaller than) the Aladin variance. For surface pressure (bottom left panel of figure H.10), this would make even the large scale Arpège variance larger than the large scale Aladin variance.

The implications of these different scenarios on a possible Aladin analysis are discussed in the next section.

### H.5.2 Implied weights in an Aladin analysis

In order to determine the initial state of Aladin, one may consider to use the Arpège analysis, the Aladin background and some observations. These three sources of information can be combined through a variational formalism. The latter approach may be seen as a generalized formulation of 3D-Var, in which the Arpège analysis is introduced as an additional source of information (Gustafsson et al. 1997, Bouttier 2002). The first two sources of information may also be combined, by using a preliminary blending technique, before doing a classical 3D-Var analysis with the observations.

In this part, we will examine the implications of the ensemble variance spectra on the information extraction in this generalized 3D-Var analysis. This will be done by considering the case where one tries to combine the Arpège analysis and the Aladin background in a similar way as in the blending technique, namely without using observations.

It may be mentioned that there are other issues which are involved in this generalized 3D-Var, and which will not be studied in the present paper. For instance, ideally, when using also observations, the Arpège analysis errors and the observation errors should be uncorrelated, in order to make the analysis procedure less complicated. This means that it may be advantageous to assimilate a large amount of observations which have not been used by the Arpège analysis.

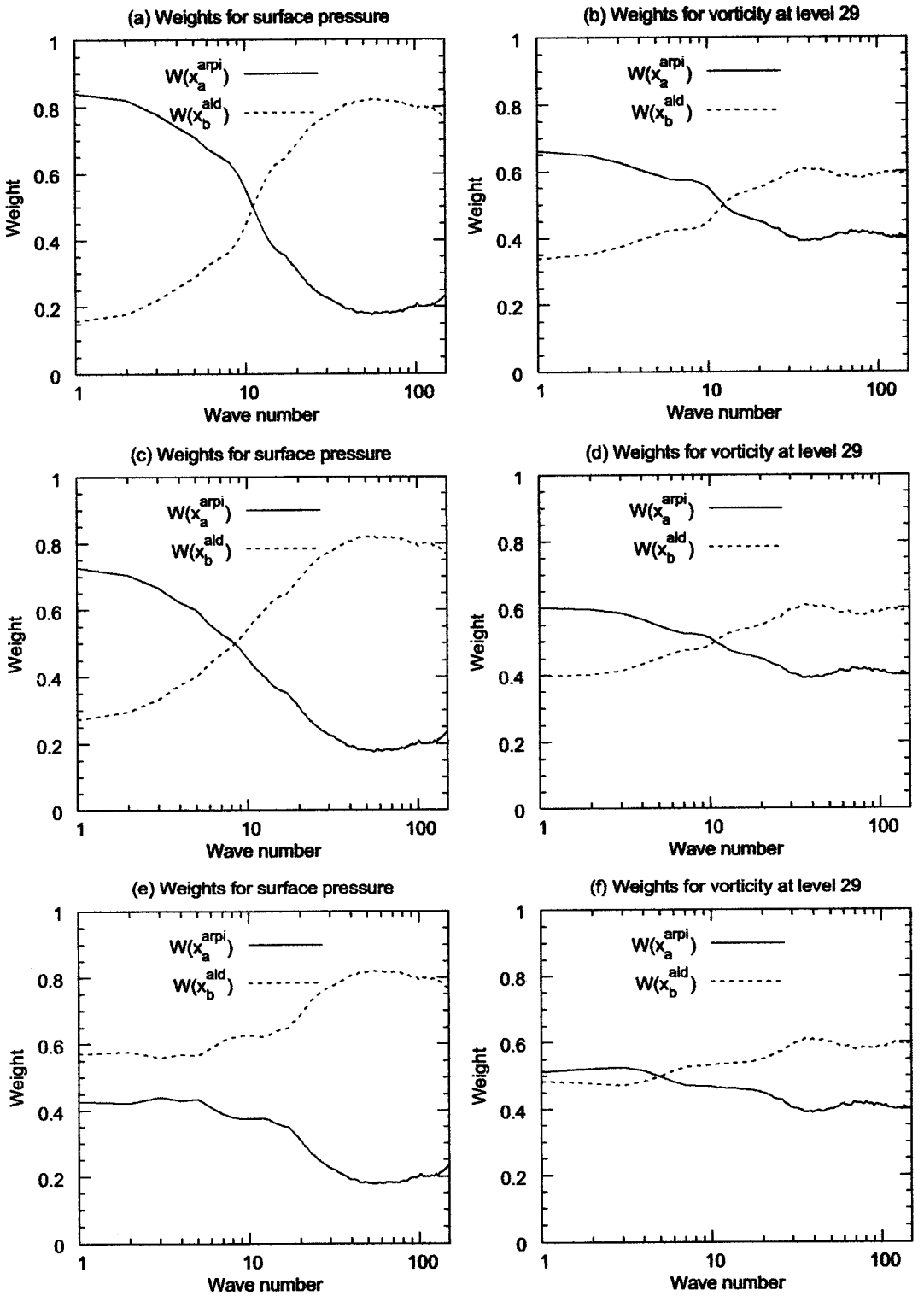


Figure H.11: The respective weights of the Arpège analysis (solid lines) and of the Aladin 6h forecast (dotted lines), resulting from their total variance spectra. Top panels: when the contribution of the large scale part of the model differences is added to the Aladin dispersion. Middle panels: when the contribution of the large scale part of the model differences is not accounted for. Bottom panels: when this large scale part is added to the Arpège dispersion.

Some examples of such observations (whose assimilation is planned for the Aladin analysis) are surface-screen level data, Meteosat-8 infrared radiances from the SEVIRI radiometer and radar data.

There may be also some correlations between the Arpège analysis errors and the Aladin background errors, for instance as the basic dynamics and physics are very similar in the two models. On the other hand, there are some model differences which arise from the Aladin high resolution and from its coupling scheme. These correlations may also depend on the amount and weight of assimilated observations in the Arpège analysis, as observation errors will contribute to the Arpège analysis errors, and as they tend to be uncorrelated with the Aladin background errors. In the remainder of the paper, we will only consider the simple case where the two involved errors are assumed to be uncorrelated.

The Arpège analysis field considered here is the initialized Arpège analysis, which is the initial state used in cold-starting mode. This implies that this Arpège field is on the same grid as the Aladin background. In other words, the equivalent of the observation operator for the Arpège field (when seeing the Arpège values as some observation values) is simply the identity matrix.

The equation that would correspond e.g. to a 3D-Var analysis $x_a^{ald}$ (without observations) is thus the following, if the two involved errors are uncorrelated (i.e., $\overline{\varepsilon_{am}^{arpi}(\varepsilon_{bm}^{ald})^T} = 0$):

$$x_a^{ald} = \mathbf{W}_a^{arpi}\, x_a^{arpi} + \mathbf{W}_b^{ald}\, x_b^{ald} \tag{H.18}$$

$\mathbf{W}_a^{arpi} = \mathbf{B}^{ald}(\mathbf{A}^{arpi} + \mathbf{B}^{ald})^{-1}$ and $\mathbf{W}_b^{ald} = \mathbf{A}^{arpi}(\mathbf{A}^{arpi} + \mathbf{B}^{ald})^{-1}$ are the respective weight matrices of the two fields, where $\mathbf{A}^{arpi}$ and $\mathbf{B}^{ald}$ are the error spatial covariance matrices of the Arpège analysis and of the Aladin background: $\mathbf{A}^{arpi} = \overline{\varepsilon_{am}^{arpi}(\varepsilon_{am}^{arpi})^T}$ and $\mathbf{B}^{ald} = \overline{\varepsilon_{bm}^{ald}(\varepsilon_{bm}^{ald})^T}$.

While the previous equations are general, in order to simplify, one can consider a univariate and 2D horizontal framework, and the case where the two covariances are homogeneous and isotropic. In this case, the two weight matrices will be diagonal in spectral space. The corresponding diagonal (scalar) values, $W_a^{arpi}$ and $W_b^{ald}$, will only depend on the 2D horizontal wave number $k^*$, and their expressions become simple functions of the ratio between the Arpège and Aladin error variances:

$$W_a^{arpi} = \frac{V(\varepsilon_{bm}^{ald})}{V(\varepsilon_{am}^{arpi}) + V(\varepsilon_{bm}^{ald})} = \frac{1}{1 + V(\varepsilon_{am}^{arpi})/V(\varepsilon_{bm}^{ald})} \tag{H.19}$$

$$W_b^{ald} = \frac{V(\varepsilon_{am}^{arpi})}{V(\varepsilon_{am}^{arpi}) + V(\varepsilon_{bm}^{ald})} = \frac{1}{1 + V(\varepsilon_{bm}^{ald})/V(\varepsilon_{am}^{arpi})} \tag{H.20}$$

The spectral variations of these weights are represented in figure H.11 for the three different scenarios. In the small scales, the weight of the Aladin forecast is the largest: its value varies from 60% (for vorticity) to 80% (for surface pressure). When attributing the large scale model differences to the Aladin coupling inaccuracies (top panels of figure H.11), it is the Arpège analysis that has the largest weight in the large scales: its maximum weight varies from around 70% (for vorticity) to around 85% (for surface pressure). The transition between these two different regimes occurs for wave numbers around 11 and 12 (which corresponds to horizontal scales of about 60-65 km), for which the two involved weights are both close to 50%.

This larger weight of Arpège in the large scales is reduced, but still present, if the term $2V(\varepsilon_{m,ls}^{ald-arpi})$ is not added to either of the two involved variances (middle panels of figure H.11): the maximum weight is around 60% for vorticity, and around 70% for surface pressure.

In the third scenario, the term $2V(\varepsilon_{m,ls}^{ald-arpi})$ is added to the Arpège variance (bottom panels of figure H.11): this implies a maximum Arpège weight which is slightly larger than 50% for vorticity, and around 40% for surface pressure.

These different figure panels thus illustrate the implications of the ensemble statistics on the information extraction in such an Aladin 3D-Var. For instance, if one of the goals (of this Aladin

analysis) is to rely on the Arpège analysis for the large scales, and on the Aladin background for the small scales, then the first two scenarios (and in particular the first one) are more appropriate for the specification of the error statistics.

## H.6 Conclusions and perspectives

An ensemble of Aladin fields has been obtained by integrating the Aladin limited area model, in cold-starting mode, from an ensemble of global Arpège analyses and forecasts. The latter global ensemble was itself obtained by integrating two independent perturbed assimilation cycles in a perfect model framework (involving perturbed observations, and perturbed backgrounds which are provided by the evolution of the previous perturbed analyses).

The evolution of the perfect model dispersion spectra has been studied, in order to examine the effects of the successive steps of an Aladin integration. The Arpège analysis reduces the large scale dispersion of the Arpège background, by using some recent observational information. The DFI reduces the small scale dispersion, as it removes some noise which is induced by the interpolation of the Arpège analysis onto the Aladin grid. Finally, the Aladin 6h forecast implies a strong increase of the small scale dispersion, in accordance with its representation of small scale dynamical and physical processes.

Then, the variances of the differences between the Aladin and Arpège 6h forecasts, that are produced from the same Arpège analysis field, have been examined. These model differences are smaller scale than the Arpège and Aladin perfect model dispersions. Moreover, it has been found that, in the small scales, the corresponding variances are very close to the difference between the variances of the Arpège and Aladin full forecast fields. This indicates that the small scale part of the model differences corresponds to some structures that are produced by the Aladin high resolution dynamics and physics, and which are not represented in Arpège (due to its lower resolution). This has allowed to propose a decomposition of the model differences variance as the sum of a small scale part which is induced directly by the resolution differences, and of a residual large scale part.

Finally, the possible implications of these model differences have been studied, for the specification of the error statistics of the Arpège analysis and of the Aladin 6h forecast. This is linked with the idea to combine these two fields (and observations) in order to provide an Aladin analysis. The variance spectra of the Arpège and Aladin dispersions indicate that the large scale information of the Aladin analysis should be extracted preferably from the Arpège analysis, which benefits from the use of some recent observations. The small scale part of the model differences variance may be added to the Arpège dispersion: this allows the representation of the larger uncertainty of the Arpège analysis with respect to the small scale features. This implies that the small scale information of the Aladin analysis would tend to be extracted from the Aladin background. The large scale part of the model differences variance is more ambiguous. It may be added to the Aladin dispersion, if it is seen as corresponding to the effects of the Aladin coupling inaccuracies for instance: this would emphasize the extraction of the large scale information from the Arpège analysis. Another possibility would be to add this large scale part to the Arpège dispersion, by considering that it reflects the beneficial (non linear) effects of the Aladin high resolution on the simulation of the large scales. However, this approach would reduce strongly the extraction of the large scale information from the Arpège analysis.

There are many natural extensions to this kind of study. Some other statistical features may be examined, such as the multivariate and three-dimensional aspects (that involve the vertical auto- and cross-covariances). Gridpoint and wavelet statistics could be computed, in order to diagnose the local (space-dependent) features of the spatial structures of the ensemble dispersion. It would also be interesting to calculate and study the cross-correlations between the Arpège and Aladin errors, and time-dependent features may be studied as well.

Moreover, an Aladin 3D-Var analysis is being developed, including a term measuring the distance to the Arpège analysis (Bouttier 2002): this will permit as well the possibility of

integrating an Aladin ensemble, in data assimilation mode. A posteriori diagnostics can then be calculated also, which may allow to compare the ensemble-based statistics with the estimates retrieved from the observations (Sadiki and Fischer 2005). The study of the model differences may be pursued, in particular by calculating differences between two models that differ either by their resolution, or by the geometry of their integration domain (global versus limited area). Another natural perspective is to assess the impact of e.g. model error modifications of the statistics on Aladin/Arpège forecasts.

*Acknowledgments.* The authors would like to thank François Bouttier, François Bouyssel, Claude Fischer and Alex Deckmyn for carefully reading the early versions of the manuscript. We would like also to thank the two reviewers for their constructive comments. Simona Ecaterina Ştefănescu and Margarida Belo Pereira are grateful to Météo France for supporting their stays at CNRM/GMAP.

## APPENDIX A

## $cov(x_b^{arpi}, \varepsilon_m^{ald-arpi})$ and $cov(x_b^{ald}, \varepsilon_m^{ald-arpi})$

*a. The covariance* $cov(x_b^{arpi}, \varepsilon_m^{ald-arpi})$

By the definition of $\varepsilon_m^{ald-arpi}$, the Aladin forecast field $x_b^{ald}$ can be written as follows: $x_b^{ald} = x_b^{arpi} + \varepsilon_m^{ald-arpi}$, which implies the following expression for the variance of the full Aladin forecast field: $V(x_b^{ald}) = V(x_b^{arpi}) + V(\varepsilon_m^{ald-arpi}) + 2\ cov(x_b^{arpi}, \varepsilon_m^{ald-arpi})$. This is equivalent to:

$$cov(x_b^{arpi}, \varepsilon_m^{ald-arpi}) = \tfrac{1}{2}[\ (V(x_b^{ald}) - V(x_b^{arpi}))\ -\ V(\varepsilon_m^{ald-arpi})\ ]$$

Therefore, the result $V(\varepsilon_m^{ald-arpi}) \sim V(x_b^{ald}) - V(x_b^{arpi})$ implies that:

$$cov(x_b^{arpi}, \varepsilon_m^{ald-arpi}) \sim 0$$

*b. The covariance* $cov(x_b^{ald}, \varepsilon_m^{ald-arpi})$

By the definition of $\varepsilon_m^{ald-arpi}$, the Arpège forecast field $x_b^{arpi}$ can be written as follows: $x_b^{arpi} = x_b^{ald} - \varepsilon_m^{ald-arpi}$, which implies the following expression for the variance of the full Arpège forecast field: $V(x_b^{arpi}) = V(x_b^{ald}) + V(\varepsilon_m^{ald-arpi}) - 2\ cov(x_b^{ald}, \varepsilon_m^{ald-arpi})$. This is equivalent to:

$$cov(x_b^{ald}, \varepsilon_m^{ald-arpi}) = \tfrac{1}{2}[\ (V(x_b^{ald}) - V(x_b^{arpi}))\ +\ V(\varepsilon_m^{ald-arpi})\ ]$$

Therefore, the result $V(\varepsilon_m^{ald-arpi}) \sim V(x_b^{ald}) - V(x_b^{arpi})$ implies that:

$$cov(x_b^{ald}, \varepsilon_m^{ald-arpi}) \sim V(\varepsilon_m^{ald-arpi})$$

which tends to be much larger than $cov(x_b^{arpi}, \varepsilon_m^{ald-arpi}) \sim 0$.

## APPENDIX B

## The covariance of perturbation differences and of single perturbations

*a. The covariance of perturbation differences*

In the two-member ensemble experiment, two independent sets of observation perturbations $\delta_o(k)$ and $\delta_o(l)$ are added to the unperturbed observation set $y$, for members $k$ and $l$ respectively: $y(k) = y + \delta_o(k)$ and $y(l) = y + \delta_o(l)$. In practice, the two sets of observation perturbations are obtained as random realizations of the Gaussian probability distribution function whose mean is zero, and whose covariance matrix $\mathbf{R}_\delta$ corresponds to the specified observation error covariance matrix $\mathbf{R}$: $\overline{\delta_o(k)(\delta_o(k))^T} = \overline{\delta_o(l)(\delta_o(l))^T} = \mathbf{R}_\delta$ with $\mathbf{R}_\delta = \mathbf{R}$. These two random realizations are uncorrelated, which implies that the covariance matrix of the observation difference $\varepsilon_o = \delta_o(k) - \delta_o(l)$ is equal to twice the covariance matrix of the observation perturbations: $\mathbf{R}_\epsilon = \varepsilon_o(\varepsilon_o)^T = \overline{\delta_o(k)(\delta_o(k))^T} + \overline{\delta_o(l)(\delta_o(l))^T} - \overline{\delta_o(k)(\delta_o(l))^T} - \overline{\delta_o(l)(\delta_o(k))^T} = 2\ \mathbf{R}_\delta$. Similarly, the analysis difference $\varepsilon_a^{arpi}$ can be seen as the difference between two independent analysis perturbations $\delta_a^{arpi}(k)$ and $\delta_a^{arpi}(l)$: $\varepsilon_a^{arpi} = \delta_a^{arpi}(k) - \delta_a^{arpi}(l)$. This implies that the covariance matrix $\mathbf{A}_\varepsilon$ of the analysis difference will be equal to twice the covariance matrix $\mathbf{A}_\delta$ of the analysis perturbations: $\mathbf{A}_\varepsilon = 2\mathbf{A}_\delta$.

*b. The covariance of single perturbations*

In sections 4 and 5, $\varepsilon_m^{ald-arpi}$ are the differences between an Aladin forecast and an Arpège forecast which are started from the same Arpège analysis. These differences arise from errors that are present in one model and not in the other one, such as the Aladin coupling errors, or the lack of representation of some small scale structures by the Arpège model (due to its low resolution). Said differently, either the Arpège model can be seen as a perturbation of the Aladin model (e.g., with an excessive dissipation of small scale structures), or the Aladin model can be seen as a perturbation of the Arpège model (e.g., with some distortions of Arpège large scale fields by the Aladin coupling procedure). In both cases, the difference $\varepsilon_m^{ald-arpi}$ is to be seen as a single realization of the probability distribution function of a model perturbation $\delta_m$

(related either to Arpège or to Aladin): $\varepsilon_m^{ald-arpi} = \delta_m$, and thus $V(\varepsilon_m^{ald-arpi}) = V(\delta_m)$ instead of $2V(\delta_m)$. This implies that a factor 2 has to be applied to $V(\varepsilon_m^{ald-arpi})$ (or to its subparts), when adding the initial condition and model error variance contributions. This allows to recover a similar weight as for the initial condition error variance.

## REFERENCES


Belo Pereira, M., L. Berre and G. Desroziers, 2002: Estimation et étude des covariances d'erreur de prévision d'ARPEGE/ALADIN (Estimation and study of ARPEGE/ALADIN background error covariances). *Proc. Ateliers de Modélisation de l'Atmosphère (AMA).* Météo France, Toulouse, 17-19 December 2002, 17-20. [Available from Météo France, 42. Av. G. Coriolis, 31057 Toulouse Cedex 01, France.]

Belo Pereira, M., and L. Berre, 2006: The use of an ensemble approach to study the background error covariances in a global NWP model. Accepted in *Mon. Wea. Rev.*

Bouttier, F., 2002: Strategies for kilometric-scale LAM data assimilation. *Report of HIRLAM workshop on variational data assimilation and remote sensing.* Finnish Meteorological Institute, Helsinki, 21-23 January 2002, 9-17. [Available from SMHI, S-60176 Norrköping, Sweden.]

Brožkova, R., D. Klaric, S. Ivatek-Šahdan, J.-F. Geleyn, V. Cassè, M. Široká, G. Radnóti, M. Janoušek, K. Stadlbacher, and H. Seidl, 2001: DFI blending: an Alternative Tool for Preparation of the Initial Conditions for LAM. *PWPR Rep. Series*, **31**, *WMO TD*, **1064**, 1-7.

Bubnová, R., A. Horányi, and S. Malardel, 1993: International project ARPEGE/ALADIN. *EWGLAM Newsletter*, **22**, Institut Royal Meteorologique de Belgique, 117-130.

Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting. *Quart. J. Roy. Meteor. Soc.*, **131**, 1013-1043.

Cohn, S. E., and D. P. Dee, 1988: Observability of discretized partial differential equations. *Siam. J. Numer. Anal.*, Vol. **25**, No. **3**, 586-617.

Courtier, P., and J.-F. Geleyn, 1988: A global numerical weather prediction model with variable resolution: Application to the shallow water equations. *Quart. J. Roy. Meteor. Soc.*, **114**, 1321-1346.

Courtier, P., C. Freydier, J.-F. Geleyn, F. Rabier, and M. Rochas, 1991: The ARPEGE project at Météo France. *Proc. ECMWF Seminar, 9-13 September 1991*, Vol. **II**, 193-231.

Daley, R., 1991: *Atmospheric data analysis.* Cambridge University Press, Cambridge, UK, 460 pp.

Daley, R., and R. Ménard, 1993: Spectral characteristics of Kalman filter systems for atmospheric data assimilation. *Mon. Wea. Rev.*, **121**, 1554-1565.

Davies, H. C., 1976: A lateral boundary formulation for multi-level prediction models. *Quart. J. Roy. Meteor. Soc.*, **102**, 405-418.

Dee, D. P., 1995: On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Wea. Rev.*, **123**, 1128-1145.

Errico, R. M., D. J. Stensrud, and K. D. Raeder, 2001: Estimation of the error distributions of precipitation produced by convective parametrization schemes. *Quart. J. Roy. Meteor. Soc.*, **127**, 2495-2512.

Fisher, M., 2003: Background error covariance modelling. *Proc. ECMWF Seminar on Recent developments in data assimilation for atmosphere and ocean, 8-12 September 2003*, 45-63.

Gustafsson, N., P. Lönnberg, and J. Pailleux, 1997: Data assimilation for high resolution limited area models. *J. Meteor. Soc. Japan*, **75**, 367-382.

Houtekamer, P. L., L. Lefaivre, J. Derome, H. Ritchie, and H. L. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, **124**, 1225-1242.

Lynch, P., and X.-Y. Huang, 1992: Initialization of the HIRLAM Model Using a Digital Filter. *Mon. Wea. Rev.*, **120**, 1019-1034.

Lynch, P., D. Giard, and V. Ivanovici, 1997: Improving the Efficiency of a Digital Filtering Scheme for Diabatic Initialization. *Mon. Wea. Rev.*, **125**, 1976-1982.

Radnóti, G., R. Ajjaji, R. Bubnová, M. Caian, E. Cordoneanu, K. von der Emde, J.-D. Gril, J. Hoffman, A. Horányi, S. Issara, V. Ivanovici, M. Janoušek, A. Joly, P. LeMoigne, and S. Malardel, 1995: The spectral limited area model ARPEGE-ALADIN. *PWPR Rep. Series*, **7**, *WMO TD*, **699**, 111-118.

Sadiki, W., C. Fischer, and J.-F. Geleyn, 2000: Mesoscale background error covariances: recent results obtained with the limited area model ALADIN over Morocco. *Mon. Wea. Rev.*, **128**, 3927-3935.

Sadiki, W., and C. Fischer, 2005: A posteriori validation applied to the 3D-Var Arpège and Aladin data assimilations system. *Tellus*, **57A**, 21-34.

Veersé, F., and J.-N. Thépaut, 1998: Multiple-truncation incremental approach for four-dimensional variational data assimilation. *Quart. J. Roy. Meteor. Soc.*, **124**, 1889-1908.

# Bibliography


[1] Andersson, E., and M. Fisher, 1998: Background errors for observed quantities and their propagation in time. *Proceedings of the ECMWF workshop on "Diagnosis of data assimilation systems"*, Reading, UK, 1-4 November 1998, 81-89.

[2] Belo Pereira, M. and L. Berre, 2005: The use of an Ensemble approach to study the Background Error Covariances in a Global NWP model. To appear in *Mon. Wea. Rev.*

[3] Berre, L., 2000: Estimation of synoptic and mesoscale forecast error covariances in a limited area model. *Mon. Wea. Rev.*, **128**, 644-667.

[4] Berre, L., S.E., Ştefănescu, and M. Belo Pereira, 2005: The representation of the analysis effect in three error simulation techniques. To appear in *Tellus*.

[5] Boer, G.J., 1983: Homogeneous and isotropic turbulence on the sphere. *J. Atmos. Sci.*, **40**, 154-163.

[6] Bougeault, P., 1985: A simple parameterisation of the large scale effects of cumulus convection. *Mon. Wea. Rev.*, **112**, 2108-2121.

[7] Bouttier, F., 1993: The dynamics of error covariances in a barotropic model. *Tellus*, **45A**, 408-423.

[8] Bouttier, F., 1994: Sur la prévision de la qualité des prévisions météorologiques. Ph. D. dissertation, Université Paul Sabatier, 240 pp. [ Available from Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex, France. ]

[9] Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F. and Fisher, M., 1998: The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part I: Formulation. *Q.J.Roy. Meteorol. Soc.*, **124**, 1783-1808.

[10] Daley, R., 1985: The analysis of synoptic scale divergence by a statistical interpolation procedure. *Mon. Wea. Rev.*, **113**, 1066-1079.

[11] Daley, R., 1991: *Atmospheric data analysis.* Cambridge University Press, Cambridge, UK, 460 pp.

[12] Deckmyn, A. and L. Berre, 2005: A wavelet approach to representing background error covariances in a limited area model. *Mon. Wea. Rev.*, **133**, 1279-1294.

[13] Derber, J., and F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195-221.

[14] Desroziers, G., V., Mathiot, and F. Orain, 1995: A study of ARPEGE forecast error covariances. *Proceedings of the WMO second international symposium on assimilation of observations in meteorology and oceanography*, Tokyo, Japan, Vol I, 263-268.

[15] Eyre, J.R., 1989a. Inversion of cloudy satellite sounding radiances by nonlinear optimal estimation. I: Theory and simulation of TOVS. *Q. J. Roy. Meteor. Soc.*, **115**, 1001-1026.

[16] Eyre, J.R., 1989b. Inversion of cloudy satellite sounding radiances by nonlinear optimal estimation. II: Application to TOVS data. *Q. J. Roy. Meteor. Soc.*, **115**, 1027-1037.

[17] Fisher, M., 2003: Background error covariance modelling. *Proceedings of the ECMWF Seminar on Recent developments in data assimilation for atmosphere and ocean, 8-12 September 2003*, 45-63.

[18] Gauthier, P., C., Charette, L., Fillion, P., Koclas, and S. Laroche, 1999: Implementation of a 3D variational data assimilation system at the Canadian Meteorological Centre. Part I: The global analysis. *Atmosphere-Ocean*, **37**, 103-156.

[19] Gustafsson, N., L., Berre, S., Hörnquist, X.-Y., Huang, M., Lindskog, B., Navascués, K. S., Mogensen, and S. Thorsteinsson, 2001: Three-dimensional variational data assimilation for a limited area model. Part I: General formulation and the background error constraint. *Tellus*, **53A**, 425-446.

[20] Hollingsworth, A. and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus* **38A**, 111-136.

[21] Hollingsworth A., 1987: Objective analysis for numerical weather prediction. Short- and medium-range numerical weather prediction. *Proc. WMO/IUGG NWP Symp.*, Tokyo, Japan, Meteorological Society of Japan, 11-59.

[22] Houtekamer, P.L., L., Lefaivre, J., Derome, H., Ritchie, and H.L. Mitchell, 1996: A System Simulation Approach to Ensemble Prediction. *Mon. Wea. Rev.*, **124**, 1225-1242.

[23] Ingleby, N.B., 2001: The statistical structure of forecast errors and its representation in The Met. Office Global 3-D Variational Data Assimilation Scheme. *Q. J. Roy. Meteor. Soc.*, **127**, 209-231.

[24] Janisková, M., J.N. Thépaut, and J.F. Geleyn, 1999: Simplified and regular physical parametrizations for incremental four-dimensional variational assimilation. *Mon. Wea. Rev.*, **127**, 26-45.

[25] Lindzen, R.S. and M. Fox-Rabinovitz, 1989: Consistent Vertical and Horizontal Resolution. *Mon. Wea. Rev.*, **117**, 2575-2583.

[26] Lönnberg, P. and A. Hollingsworth, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus* **38A**, 137-161.

[27] Lorenc, A., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, **109**, 701-721.

[28] Lorenc, A., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112**, 1177-1194.

[29] Lorenc, A.C., S.P. Ballard, R.S., Bell, N.B., Ingleby, P.L.F., Andrews, D.M., Barker, J.R., Bray, A.M., Clayton, T., Dalby, D., Li, T.J., Payne, and F.W. Saunders, 2000: The Met. Office Global 3-Dimensional Variational Data Assimilation Scheme. *Q. J. Roy. Meteor. Soc.*, **126**, 2991-3012.

[30] Louis, J.F., 1979: A parametric model of vertical eddy fluxes in atmosphere *Boundary Layer Meteorology.*, **17**, 187-202.

[31] Louis, J.F., M., Tiedtke and J.F. Geleyn, 1979: A short history of the operational PBL. *Proceedings of the ECMWF workshop on the "Planetary boundary layer parameterisation"*, Reading, UK, 25-27 November 1981, 59-80.

[32] Mahfouf, J.F. and F., Rabier, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. II: Experimental results with improved physics. *Q. J. Roy. Meteor. Soc.*, **126**, 1171-1190.

[33] Mc Nally, A., 2000: Estimates of short-term forecast-temperature error correlations and the implications for radiance-data assimilation, *Q.J.R. Meteorol. Soc.*, **126**, 361-373.

[34] McPherson, R. D., K. H. Bergman, R. E. Kistler, G.E., Rasch, and D.S. Gordon , 1979: The NMC operational global data assimilation scheme. *Mon. Wea. Rev.*, **107**, 1445-1461.

[35] Mitchell, H.L., P.L., Houtekamer, and G. Pellerin, 2002: Ensemble Size, Balance, and Model-Error Representation in an Ensemble Kalman Filter. *Mon. Wea. Rev.*, **130**, 2791-2808.

[36] Parrish, D.F. and J.C. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747-1763.

[37] Rabier, F., A., Mc Nally, E., Andersson, P., Courtier, P., Undén, J., Eyre, A., Hollingsworth, and F. Bouttier, 1998: The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part II: Structure functions. *Quart. J. Roy. Meteor. Soc.*, **124**, 1809-1829.

[38] Rabier, F., H., Järvinen, E., Klinker, J.-F., Mahfouf, and A.J. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. Part I: Experimental results with simplified physics. *Q.J.R. Meteorol. Soc.*, **126**, 1143-1170.

[39] Ştefănescu S. E., L., Berre and M. Belo Pereira, 2005: The evolution of dispersion spectra and the evaluation of model differences in an ensemble estimation of error statistics for a limited area model analysis. Accepted to *Mon. Wea. Rev.*.

[40] Thepaut J. N., R. N. Hoffman, and P. Courttier, 1993: Interactions of dynamics and observations in a 4-dimensional variational assimilation. *Mon. Wea. Rev.*, **121**, 3393-3414.

[41] Undén, P., 1989: Tropical Data Assimilation and Analysis of Divergence. *Mon. Wea. Rev.*, **117**, 2495-2517.

[42] Veersé, F. and J.-N. Thépaut, 1998: Multiple-truncation incremental approach for four-dimensional variational data assimilation. *Q.J.R. Meteorol. Soc.*, **124**, 1889-1908.

[43] Wu, W.-S., R., Purser, and D. Parrish, 2002: Three-Dimensional Variational Analysis with Spatially Inhomogeneous Covariances. *Mon. Wea. Rev.*, **130**, 2905-2916.

[44] Žagar N., E., Andersson, and M. Fisher, 2004: Balanced tropical data assimilation based on a study of equatorial waves in ECMWF short-range forecast errors. *ECMWF Technical Memorandum* No. 437, 30pp.

# List of Figures

# List of Tables